Outer Circles

An Introduction to Hyperbolic 3-Manifolds

Albert Marden

CAMBRIDGE

CAMBRIDGE www.cambridge.org/9780521839747

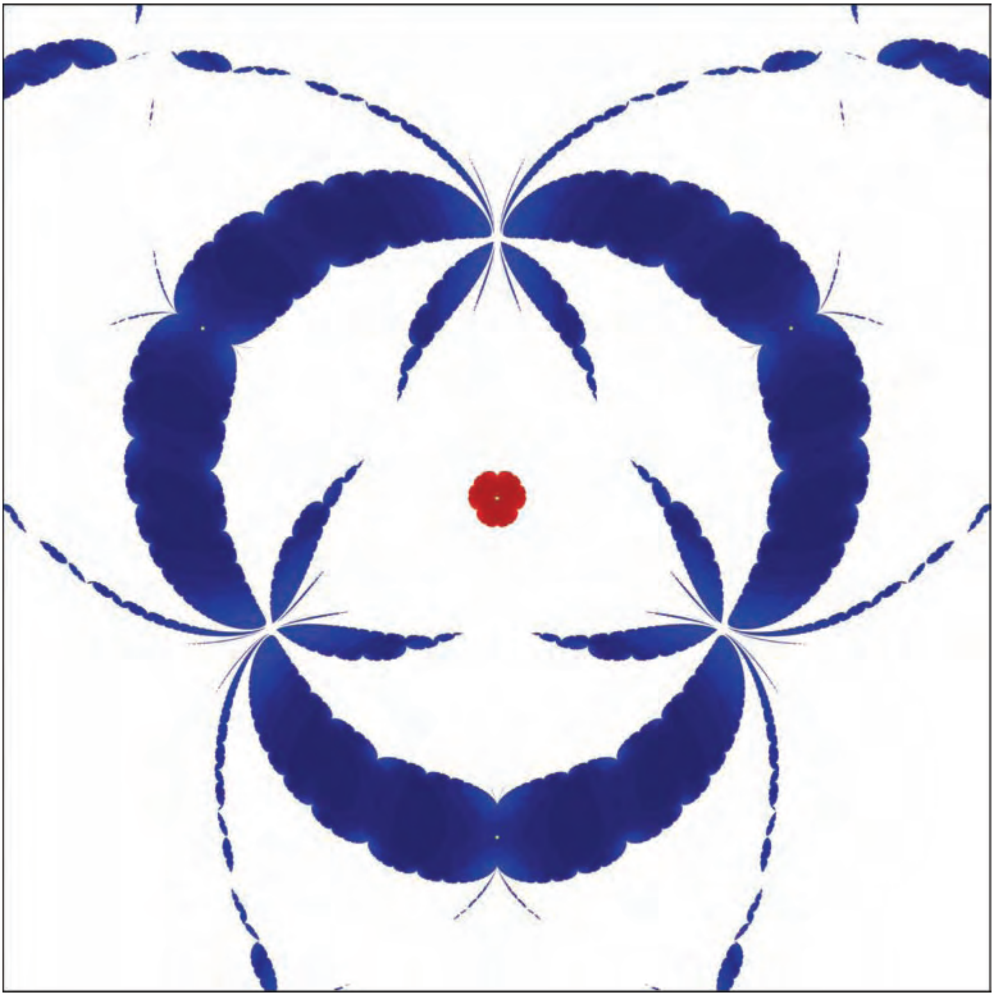
This page intentionally left blank

OUTER CIRCLES

We live in a three-dimensional space; what sort of space is it? Can we build it from simple geometric objects? The answers to such questions have been found in the last 30 years, and *Outer Circles* describes the basic mathematics needed for those answers as well as making clear the grand design of the subject of hyperbolic manifolds as a whole.

The purpose of *Outer Circles* is to provide an account of the contemporary theory, accessible to those with minimal formal background in topology, hyperbolic geometry, and complex analysis. The text explains what is needed, and provides the expertise to use the primary tools to arrive at a thorough understanding of the big picture. This picture is further filled out by numerous exercises and expositions at the ends of the chapters and is complemented by a profusion of high quality illustrations. There is an extensive bibliography for further study.

ALBERT MARDEN is a Professor of Mathematics in the School of Mathematics at the University on Minnesota.



The discreteness locus in the extended Bers slice of the hexagonal once-punctured torus (see Exercise 6-8). The Bers slice—the red central object—is surrounded by other islands of discontinuity, in blue. The inward pointing cusps on the Bers slice boundary represent geometrically finite groups and the same is presumably true for the other components. The yellow dots are the fuchsian centers of the components. Only a small number of islands are shown because of theoretical and computational limitations.

The computation and image were made by David Dumas of Brown University; his web site contains many beautiful related images.

OUTER CIRCLES

An Introduction to Hyperbolic 3-Manifolds

A. MARDEN
University of Minnesota

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK
Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521839747

First published in print format 2007

ISBN-13 978-0-511-28901-9 eBook (EBL)
ISBN-10 0-511-28901-4 eBook (EBL)

ISBN-13 978-0-521-83974-7 hardback
ISBN-10 0-521-83974-2 hardback

To Dorothy

צו מיין פרוי דבורה, די אמת'דיקע אשת-חיל.

Contents

List of Illustrations

Credits: CM, Curt McMullen; CS, Caroline Series; DD, David Dumas; DW, Dave Wright; HP, Howard Penner; JB, Jeff Brock; JP, John Parker; KS, Ken Stephenson; RB, Robert Brooks; SL, Silvio Levy; YM, Yair Minsky.

Preface

To a topologist a teacup is the same as a bagel, but they are not the same to a geometer. By analogy, it is one thing to know the topology of a 3-manifold, another thing entirely to know its geometry — to find its shortest curves and their lengths, to make constructions with polyhedra, etc. In a word, we want to do geometry in the manifold just like we do geometry in euclidean space.

But do general 3-manifolds have "natural" metrics? For a start we might wonder when they carry one of the standards: the euclidean, spherical or hyperbolic metric. The latter is least known and not often taught; in the stream of mathematics it has always been something of an outlier. However it turns out that it is a big mistake to just ignore it! We now know that the interior of " most" compact 3-manifolds carry a hyperbolic metric.

It is the purpose of this book to explain the geometry of hyperbolic manifolds. We will examine both the existence theory and the structure theory.

Why embark on such a study? Well after all, we do live in three dimensions; our brains are specifically wired to see well in space. It seems perfectly reasonable if not compelling to respond to the challenge of understanding the range of possibilities. In particular, it is not at all established that our own universe is euclidean space, as many so like to believe.

I will briefly summarize the recent history of our subject. Although Poincaré recognized in 1881 that Möbius transformations extend from the complex plane to upper half-space, the development of the theory of three-dimensional hyperbolic manifolds had to wait for progress in three-dimensional topology. It was as late as the mid-1950s that Papakyriakopoulos confirmed the validity of Dehn's Lemma and the Loop Theorem. Once that occurred, the wraps were off.

In the early 1960s, while 3-manifold topology was booming ahead, the theory of kleinian groups was abruptly awoken from its long somnolence by a brilliant discovery of Lars Ahlfors. Kleinian groups are the discrete isometry groups of hyperbolic 3-space. Working (as always) in the context of complex analysis, Ahlfors discovered their finiteness property. This was followed by Mostow's contrasting discovery that closed hyperbolic manifolds of dimension $n \geq 3$ are uniquely determined up to isometry by their isomorphism class. This too came as a bombshell as it is false for $n = 2$. Then came Bers' study of quasifuchsian groups and his and Maskit's fundamental discoveries of "degenerate groups" as limits of them. Along a different line, Jørgensen developed the methods for dealing with sequences of kleinian groups, recognizing the existence of two distinct kinds of convergence which he called "algebraic" and "geometric". He also discovered a key class of examples, namely hyperbolic 3-manifolds that fiber over the circle.

It wasn't until the late 1960s that 3-manifold topology was sufficiently understood, most directly by Waldhausen's work, and the fateful marriage of 3-manifold topology to the complex analysis of the group action on $\mathbb{S}^2$ occurred. The first application was to the classification and analysis of geometrically finite groups and their quotient manifolds.

During the 1960s and 1970s, Riley discovered a slew of faithful representations of knot and link groups in $\mathrm{PSL}(2,\mathbb{C})$. Although these were seen as curiosities at the time, his examples pressed further the question of just what class of 3-manifolds did the hyperbolic manifolds represent? Maskit had proposed using his combination theorems to construct all hyperbolic manifolds from elementary ones. Yet Peter Scott pointed out that the combinations that were then feasible would construct only a limited class of 3-manifolds.

So by the mid-1970s there was a nice theory, part complex analysis, part three-dimensional geometry and topology, part algebra. Noone had the slightest idea as to what the scope of the theory really was. Did kleinian groups represent a large class of manifolds, or only a small sporadic class?

The stage (but not the players) was ready for the dramatic entrance in the mid-1970s of Thurston. He arrived with a proof that the interior of "most" compact 3-manifolds has a hyperbolic structure. He brought with him an amazingly original, exotic, and very powerful set of topological/geometrical tools for exploring hyperbolic manifolds. The subject of two- and three-dimensional topology and geometry was never to be the same again.

This book. Having witnessed at first hand the transition from a special topic in complex analysis to a subject of broad significance and application in mathematics, it seemed appropriate to write a book to record and explain the transformation. My idea was to try to make the subject accessible to beginning graduate students with minimal specific prerequisites. Yet I wanted to leave students with more than a routine compendium of elementary facts. Rather I thought students should see the big picture, as if climbing a watchtower to overlook the forest. Each student should end his or her studies having a personal response to the timeless question: What is this good for?

With such thoughts in mind, I have tried to give a solid introduction and at the same time to provide a broad overview of the subject as it is today. In fact today, the subject has reached a certain maturity. The characterization those compact manifolds whose interiors carry a hyperbolic structure is complete, the final step being provided by Perelman's recent confirmation of the Geometrization Conjecture. Attention turned to the analysis of structure of hyperbolic manifolds assuming only a finitely generated fundamental group. Within the past few years, the structure of these has been worked out as well. The three big conjectures left over from the 1960s and 1970s have been solved: tameness, density, and classification of the ends (ideal boundary components). If one is willing to climb the watchtower, the view is quite remarkable.

It is a challenge to carry out the plan as outlined. The foundation of the subject rests on elements of three-dimensional topology, hyperbolic geometry, and modern complex analysis. None of these are regularly covered in courses at most places.

I have attempted to meet the challenge as follows. The presentation of the basic facts is fairly rigorous. These are included in the first four chapters, plus the optional Chapters 7 and 8. These chapters include crash courses in three-manifold topology, covering surfaces and manifolds, quasiconformal mappings, and Riemann surface theory. With the basic information under our belts, Chapters 5 and 6 (as well as parts of Chapters 3 and 4) are expository, without most proofs. The reader will find there both the Hyperbolization Theorem and the newly discovered structural properties of general hyperbolic manifolds.

At the end of each chapter is a long section titled "Exercises and Explorations". Some of these are genuine exercises and/or important additional information directly related to the material in the chapter. Others dig away a bit at the proofs of some of the theorems by introducing new tools they have required. Still others are included to point out various paths one can follow into the deeper forest and beauty spots one can find there. Thus there are not only capsule introductions to big fields like geometric group theory, but presentations of other more circumscribed topics that I (at least) find fascinating and relevant.

Acknowledgments. It is a great pleasure to thank the people who have helped bring the book to fruition.

First I want to acknowledge the essential contributions of my friend and colleague Troels Jørgensen. Over more than 25 years we walked in the forest together discussing and admiring the landscape our studies revealed. In particular we discussed the "universal properties" of Chapter 3 for years, until it was too late to publish them. Chapters 7 and 8 are based on his private lectures.

David Wright kindly computed a number of limit sets of kleinian groups, some never before seen, others adapted from pictures created for *Indra's Pearls* [Mumford et al. 2002]. The extent of his contribution is evident from the list of figures. His pictures can be downloaded from www.okstate.edu/~wrightd/Marden together with computational details. In addition, David Dumas was willing to share his visualization of a Bers slice amidst the surrounding archipelago of discreteness components. It serves as the frontispiece. Jeff Brock contributed his pictures of algebraic and geometric limits that originally appeared in [Brock 2001b]; these too can be seen on `www.math.brown.edu/~brock`. The presence of the many artfully crafted pictures is a tangible expression of the mathematical beauty of the subject.

I am very grateful to Ken'ichi Ohshika for reading and commenting on an early draft and Dick Canary for reading several chapters of a later draft of the manuscript. Ian Agol, Ken Bromberg, Richard Evans, Sadayoshi Kojima, Howie Masur, Vlad Markovic, Yair Minsky, Peter Scott and Juan Souto as well as other mathematicians have been generous in responding to specific questions as well.

I could not have completed the book in the present form without the expert guidance and participation of Silvio Levy. He identified math problems, fixing some of them, properly handled the LaTeX formatting, improved the syntax, crafted the diagrams, and inserted the pictures.

I want to acknowledge the institutional support from the Forschungsinstitut für Mathematik at ETH in Zurich, the Maths Research Center, University of Warwick, and not least, from my own department, the School of Mathematics of the University of Minnesota. In my semester course Math 8380, I was able to present a solid introduction and overview of the subject based on the main points in the first six chapters.

I am grateful to Caroline Series for introducing me to the Press and for her enthusiasm for the project. Cambridge University Press in the person of David Tranah has shown great flexibility in keeping the retail price down and publishing standards high. Most importantly, David provided Silvio Levy as editor.

The nineteenth-century history. The history of noneuclidean geometry in the early nineteenth century is fascinating because of a host of conflicted issues concerning axiom systems in geometry, and the nature of physical space [Gray 1986; 2002].

Jeremy Gray [2002] writes:

> Few topics are as elusive in the history of mathematics as Gauss's claim to be a, or even the, discoverer of Non-Euclidean geometry. Answers to this conundrum often depend on unspoken, even shifting, ideas about what it could mean to make such a discovery. ... [A]mbiguities in the theory of Fourier series can be productive in a way that a flawed presentation of a new geometry cannot be, because there is no instinctive set of judgments either way in the first case, but all manner of training, education, philosophy and belief stacked against the novelties in the second case.

Gray goes on to quote from Gauss's 1824 writings:

> ...the assumption that the angle sum is less than 180° leads to a geometry quite different from Euclid's, logically coherent, and one that I am entirely satisfied with. It depends on a constant, which is not given a priori. The larger the constant, the closer the geometry to Euclid's.... The theorems are paradoxical but not self-contradictory or illogical.... All my efforts to find a contradiction have failed, the only thing that our understanding finds contradictory is that, if the geometry were to be true, there would be an absolute (if unknown to us) measure of length a priori. ... As a joke I've even wished Euclidean geometry was not true, for then we would have an absolute measure of length a priori.

From his detailed study of the history, Gray's conclusion expressed in his recent Zurich lecture is that the birth of noneuclidean geometry should be attributed to the independently written foundational papers of Lobachevsky in 1829 and Bólyai in 1832. As expressed in [Milnor 1994, p. 246], those two were the first "with the courage to publish" accounts of the new theory. Still,

> [f]or the first forty years or so of its history, the field of non-euclidean geometry existed in a kind of limbo, divorced from the rest of mathematics, and without any firm foundation.

This state of affairs changed upon Beltrami's introduction in 1868 of the methods of differential geometry, working with constant curvature surfaces in general. He gave the first global description of what we now call hyperbolic space. See [Gray 1986, p. 351], [Milnor 1994, p. 246], [Stillwell 1996, pp. 7–62].

It was Poincaré who brought two-dimensional hyperbolic geometry into the form we study today. He showed how it was relevant to topology, differential equations, and number theory. Again I quote Gray, in his translation of Poincaré's work of 1880 [Gray 1986, p. 268–9].

> There is a direct connection between the preceding considerations and the non-Euclidean geometry of Lobachevskii. What indeed is a geometry? It is the study of a *group of operations* formed by the displacements one can apply to a figure without deforming it. In Euclidean geometry this group reduces to *rotations* and *translations*. In the pseudo-geometry of Lobachevskii it is more complicated... [Poincaré's emphasis].

As already mentioned, the first appearance of what we now call Poincaré's conformal model of noneuclidean space was in his seminal 1881 paper on kleinian groups. He showed that the action of Möbius transformations in the plane had a natural extension to a conformal action in the upper half-space model.

Actually the names "fuchsian" and "kleinian" for the isometry groups of two- and three-dimensional space were attached by Poincaré. However Poincaré's choice more reflects his generosity of spirit toward Fuchs and Klein than the mathematical reality. Klein himself objected to the name "fuchsian". His objection in turn prompted Poincaré to introduce the name "kleinian" for the discontinuous groups that do not preserve a circle. The more apt name would perhaps have been "Poincaré groups" to cover both cases. For the full story see [Gray 1986, §6.4].

So here we are today, nearly 125 years after Poincaré and approaching 200 after the initial ferment of ideas of Gauss, witnessing a full flowering of the vision and struggle for understanding of the nineteenth-century masters.

Albert Marden
am@umn.edu
Minneapolis, Minnesota
May 19, 2006

1

Hyperbolic space and its isometries

In this chapter we gather together basic information about the geometry of two- and three-dimensional hyperbolic spaces and their isometries. This will set the stage for our study of quotient manifolds and orbifolds which begins in the next chapter.

1.1 Möbius transformations

A *Möbius transformation* in the unit sphere $\mathbb{S}^n$ of dimension n is, by definition, the result of a composition of reflections in $(n-1)$-dimensional spheres in $\mathbb{S}^n$. It will be orientation preserving if it is the composition of an even number of reflections. A defining property is that Möbius transformations send $(n-1)$-dimensional spheres onto $(n-1)$-dimensional spheres. Automatically, a symmetric pair of points (with respect to reflection) about one sphere gets sent to a symmetric pair about the other.

From now on, the unqualified term ***Möbius transformation*** *will be reserved for those that preserve orientation.* The orientation reversing kind will be called anti-Möbius transformations. For a discussion of the latter, see Exercise 1-31 at the end of the chapter.

The study of hyperbolic 3-manifolds is intimately connected with the study of Möbius and anti-Möbius transformations on the two-dimensional sphere $\mathbb{S}^2$. Via

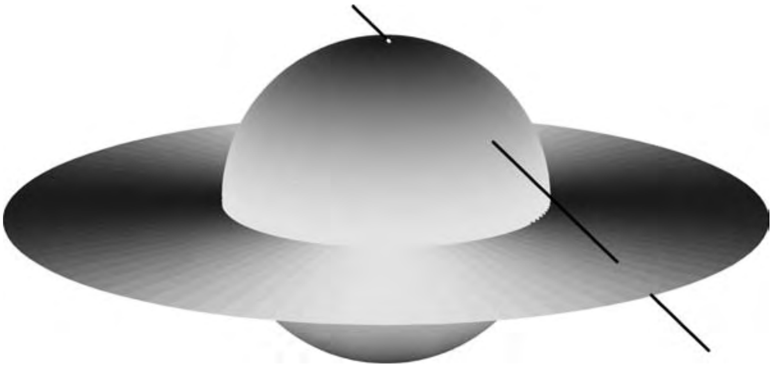


Fig. 1.1. Stereographic projection

stereographic projection (Figure 1.1), $\mathbb{S}^2$ is homeomorphic to the extended plane $\mathbb{C}\cup\infty$, and we will freely use this fact to change points of view between the extended plane and the 2-sphere. Under stereographic projection, the collection of circles and straight lines in $\mathbb{C}$ corresponds to the collection of circles on $\mathbb{S}^2$; a straight line in $\mathbb{C}$ corresponds to a circle on $\mathbb{S}^2$ through the north pole. With this correspondence in mind, we can refer to the collection of circles and lines in $\mathbb{C}$ simply as "circles". Moreover stereographic projection is a conformal map, that is, it preserves angles between intersecting arcs — in particular, angles of intersection between circles.

Möbius transformations in two dimensions are *fractional linear transformations* of the extended plane. That is, a Möbius transformation acting on $\mathbb{C}\cup\infty$ has the form

$$z \mapsto A(z) = \frac{az+b}{cz+d}, \quad \text{with } a, b, c, d \in \mathbb{C} \text{ such that } ad-bc \neq 0. \tag{1.1}$$

(When $ad-bc=0$ the expression on the right is a constant, so the map is not a Möbius transformation.) As we will see shortly, a map of this form can indeed be expressed as the composition of an even number of reflections in circles (in fact, two or four circles: see Exercise 1-7). The symmetry properties of such maps are established in Exercise 1-2.

Möbius transformations are conformal maps. In fact, the only conformal homeomorphisms of $\mathbb{C}\cup\infty$ are Möbius transformations.

We will generally assume that the representation in (1.1) is *normalized*, meaning that $ad-bc=1$. Then we can identify the group of Möbius transformations with the quotient $\mathrm{PSL}(2,\mathbb{C}) := \mathrm{SL}(2,\mathbb{C})/\pm I$, where $\mathrm{SL}(2,\mathbb{C})$ is the group of 2×2 matrices of determinant one and I is the identity matrix:

$$A(z) = \frac{az+b}{cz+d} \longleftrightarrow \pm\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad A^{-1}(z) \longleftrightarrow \pm\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

The $\pm$ ambiguity cannot be avoided. We will not keep inserting it, unless it plays an essential role. In any case the value of changing from transformations to matrices lies mainly in the algebra of composition. If A, B are Möbius transformations, the Möbius transformation resulting from the application of A followed by B is written BA; the corresponding matrix is just the usual product BA of the component matrices, in the order written. The $\pm$ ambiguity follows along. We will hop from one to the other, the representation as a transformation to the representation as a matrix, depending on which best suits the situation, without changing the labeling.

Two Möbius transformations A, B are *conjugate* if there is a Möbius transformation U such that $B = UAU^{-1}$. Conjugate transformations have the same geometry: U effects transfer of the geometry of A to that of B.

The expression $ABA^{-1}B^{-1}$ is called the *commutator* of A and B and written as $[A, B]$. Two elements commute if and only if their commutator is the identity. *

* The alternative conventions $[A, B] = B^{-1}A^{-1}BA$ or $A^{-1}B^{-1}AB$ are preferred by some authors; they do the same job, but the formulas come out differently.

The *trace* of a Möbius transformation A is, by definition, the trace of the normalized matrix of A:

$$\tau_A = \operatorname{tr} A = \pm(a+d).$$

It is invariant under conjugation. The $\pm$ ambiguity can be avoided either by using τ_A^2 or by specifying $0 \le \arg \tau_A < \pi$.

By solving the equation $A(z) = z$, we find that a nontrivial Möbius transformation has one or two fixed points in $\mathbb{S}^2$, namely $(a - d \pm \sqrt{\tau_A{}^2 - 4})/2c$, when $c \neq 0$, or otherwise the points ∞ and $b/(d-a) = ab/(1-a^2)$. Here $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, $ad - bc = 1$. Only the identity can have three fixed points.

Given three distinct points $(p_2, p_3, p_4) \in \mathbb{S}^2$, there exists a necessarily unique Möbius transformation sending p_2 to 1, p_3 to 0, p_4 to ∞. It is given by

$$z \mapsto \frac{(z - p_3)(p_2 - p_4)}{(z - p_4)(p_2 - p_3)} = (z, p_2, p_3, p_4),$$

when none of the points p_i is ∞. By taking the limit as some $p_i \to \infty$, we obtain the correct expression for $p_i = \infty$. The expression (z, p_2, p_3, p_4) is called the *cross ratio* of the four points. * Cross ratios are invariant under Möbius transformations:

$$(Az, Ap_2, Ap_3, Ap_4) = (z, p_2, p_3, p_4) \quad \text{for any } A.$$

This is a consequence of the fact that $T(z) = (z, p_1, p_2, p_3)$ satisfies $T \circ A^{-1}(z) = (z, Ap_1, Ap_2, Ap_3)$.

Apart from the identity, Möbius transformations fall into one of three types:

A is *parabolic* if the following equivalent properties hold.

- A is conjugate to $z \mapsto z + 1$.
- A has exactly one fixed point in $\mathbb{S}^2$.
- $\tau_A = \pm 2$ and $A \neq \mathrm{id}$.

A is *elliptic* if the following equivalent properties hold.

- A is conjugate to $z \mapsto e^{2i\theta} z$, with $2\theta \not\equiv 2\pi$.
- $\tau_A \in (-2, +2)$.
- A has exactly two fixed points, and the derivative of A has absolute value 1 at each of them.

A is *loxodromic* if the following equivalent properties hold.

- A is conjugate to $z \mapsto \lambda^2 z$, with $|\lambda| > 1$.
- $\tau_A \in \mathbb{C} \setminus [-2, +2]$.
- A has exactly two fixed points, one attracting and one repelling.

We will use the term *standard forms* for the conjugates for the conjugates just listed. The geometry of a general normalized Möbius transformation A is most easily read off from the conjugate standard form. Note that the elliptic $z \mapsto 1/z$ is conjugate to $z \mapsto -z$.

* The definition given has the property $(z, 1, 0, \infty) = z$. A common alternate definition results in $(z, 0, 1, \infty) = z$.

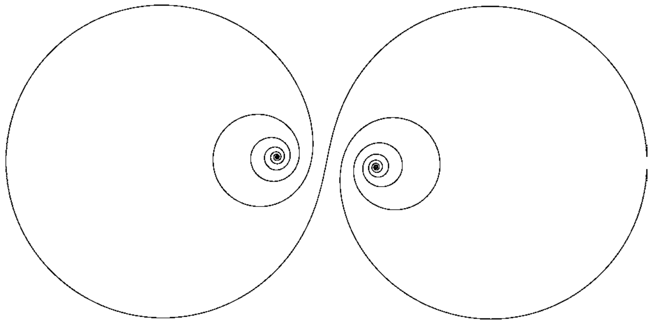


Fig. 1.2. Invariant spiral of a loxodromic with trace $\lambda+\lambda^{-1}=1.976+0.005i$.

A loxodromic Möbius transformation A has a collection of *loxodromic curves* or *invariant spirals* in $\mathbb{S}^2$. (In navigation, a *loxodromic curve* or *rhumb line* is a path of constant bearing: it makes equal oblique angles with all meridians, and so coils around the poles without ever reaching them.) For the standard form $z \mapsto \lambda^2 z$, one such spiral is given by

$$z(t)=\lambda^{2t}, \quad -\infty < t < \infty.$$

If σ denotes the segment $0 \le t < 1$ of the spiral, the various images $\{A^n(\sigma)\}$ cover the spiral without overlap. See Figure 1.2.

For additional structure in special cases see [Wright 2006].

The term *hyperbolic transformation* has historically been used to designate a loxodromic transformation whose trace is real. Such a transformation is conjugate to $z \mapsto \lambda^2 z$ with $\lambda > 1$. Nowadays the term "hyperbolic" is also used for a loxodromic element acting in hyperbolic 3-space.

The classification is proved by first conjugating A so that one fixed point lies at ∞ and the other, if there is one, at 0. The further conjugation $z \mapsto 1/z$ that interchanges 0 and ∞ may be needed to put the attracting fixed point at ∞.

If $p \in \mathbb{C}$ is a fixed point of $A \neq \mathrm{id}$, p is attracting if and only if $|A'(p)| < 1$ and repelling if and only if $|A'(p)| > 1$. The transformation A is parabolic if and only if $A'(p)=1$; A is elliptic if and only if $|A'(p)|=1$ but $A'(p) \neq 1$.

Upon referring to the normalized matrix $A=\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, we find that the *eigenvalues* are $\lambda, \lambda^{-1}=\frac{1}{2}(\operatorname{tr} A \pm \sqrt{\operatorname{tr}^2 A-4})$. The corresponding eigenvectors $\left(\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right)$ satisfy

$$\frac{\alpha}{\beta}=\frac{\lambda-d}{c}=p \quad \text{and} \quad \frac{\alpha}{\beta}=\frac{\lambda^{-1}-d}{c}=q,$$

where p, q are the fixed points. Like the trace, the eigenvalues are invariant under conjugation. The eigenvalues of an elliptic transformation have the form $e^{\pm i\theta}$ and the trace is $2\cos\theta$. A loxodromic transformation has eigenvalues $\lambda^{\pm 1}$ and trace $\lambda+\lambda^{-1}$. We can choose λ so that $|\lambda|>1$, that is, so that λ is the expanding eigenvalue.

The expanding eigenvalue of a loxodromic element A can be expressed as a cross ratio by the formula

$$\lambda^2 = (z, A(z), p_+, p_-),$$

where p_+, p_- are the attracting and repelling fixed points. (It is enough to confirm this when $p_+ = \infty$ and $p_- = 0$.)

We can write $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ as

$$Az = \frac{1}{-c^2(z + d/c)} + \frac{a}{c} \quad \text{if } c \neq 0, \qquad Az = \frac{a}{d}\left(z + \frac{b}{a}\right) \quad \text{if } c = 0. \tag{1.2}$$

This expresses A in terms of simple building blocks: maps in standard form, plus the map $z \mapsto 1/z$. Each of these has the property of preserving (generalized) circles. Therefore any Möbius transformation preserves circles, as mentioned earlier. Likewise each building block is easily seen to be a composition of two reflections, so a Möbius transformation is the composition of an even number of reflections.

Three distinct points p_2, p_3, p_4 uniquely determine a circle C, with an orientation determined by their order. When C is a proper circle, we say that the orientation thus defined is *positive* if the interior of the circle lies to the left as p_2, p_3, p_4 are encountered in that order. Let q_2, q_3, q_4 be another set of distinct points, and C' the circle through them. The Möbius transformation T that sends $p_i \to q_i$ automatically sends C onto C'. If both are proper circles, T sends the interior of C to the interior of C' if and only if the triples give both circles positive (or negative) orientations. The transformation $T : z \to w$ can be expressed in terms of cross ratios as

$$(w, q_2, q_3, q_4) = (z, p_2, p_3, p_4).$$

But if we focus simply on sending C to C', and a designated side of C to a designated side of C', it is more efficient to find T by cross ratio using the symmetry property: A Möbius transformation sends points symmetric with respect to reflection in one circle, to a pair of points symmetric in the image (Exercise 1-2). For a proper circle, the most conspicuous symmetric points are its center and ∞.

A cross ratio (p, p_2, p_3, p_4) is real if and only if the four points lie on a circle in $\mathbb{S}^2$. The cross ratio is positive if and only if (p, p_3, p_4) gives the circle the same orientation as (p_2, p_3, p_4).

We are now ready to show that Möbius transformations in $\mathbb{C} \cup \infty$ can be extended to Möbius transformations acting in upper half-space $\{\vec{x} = (z, t) : z \in \mathbb{C},\ t > 0\}$. The simplest way to see this is by applying the following observation. Each Möbius transformation is the composition of an even number of reflections in circles or lines in $\mathbb{C}$. A reflection in a circle extends naturally to the reflection in the upper hemisphere bounded by that circle. Likewise the reflection in a straight line extends to the reflection in the vertical half-plane bounded by that line. (The same argument shows that Möbius transformations on $\mathbb{S}^n = \mathbb{R}^n \cup \{\infty\}$ extend to upper half $(n+1)$-space.)

A Möbius transformation acting on $\mathbb{C} \cup \infty$ sends a given circle to another circle or line. Its extension to upper half-space will therefore map the hemisphere bounded by the circle to the hemisphere or half-plane bounded by the image of the circle. We conclude that the extension to upper half-space maps the totality of hemispheres and vertical half-planes onto itself.

If two hemispheres intersect, or a hemisphere and a vertical half-plane intersect, the intersection is a semicircle which is orthogonal to $\mathbb{C}$. If two vertical half-planes intersect, they intersect in a vertical half-line orthogonal to $\mathbb{C}$. The extension of a Möbius transformation thus maps the totality of half-lines and semicircles orthogonal to $\mathbb{C}$ onto itself. The dihedral angles between intersecting hemispheres is the same as the angle of intersection between their bounding circles in $\mathbb{C}$.

It is useful to explicitly work out the formula for extension to upper half-space $\{\vec{x} = (z, t) : z \in \mathbb{C},\ t > 0\}$. We first extend the building blocks. First,

$$\begin{aligned} z \mapsto az \quad & \text{becomes} \quad (z, t) \mapsto (az, |a|\, t); \\ z \mapsto z + b \quad & \text{becomes} \quad (z, t) \mapsto (z + b, t). \end{aligned}$$

The inversion $z \mapsto z^{-1}$ is most easily dealt with as the composition of two anti-Möbius transformations: $z \mapsto \bar{z}$ (reflection in a line) and $z \mapsto z/|z|^2 = \bar{z}^{-1}$ (reflection in the unit circle). Extending to reflections in a vertical plane and the unit hemisphere, we get respectively $(z, t) \mapsto (\bar{z}, t)$ and

$$\vec{x} \mapsto \frac{\vec{x}}{|\vec{x}|^2} \quad \text{or} \quad (z, t) \mapsto \left(\frac{z}{|z|^2 + t^2}, \frac{t}{|z|^2 + t^2} \right).$$

Therefore,

$$z \mapsto \frac{1}{z} \quad \text{becomes} \quad (z, t) \mapsto \left(\frac{\bar{z}}{|z|^2 + t^2}, \frac{t}{|z|^2 + t^2} \right).$$

Composing the building blocks we find that the extension of $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is

$$\begin{aligned} (z, t) &\mapsto \left(-\frac{\overline{z + d/c}}{c^2\big(|z + d/c|^2 + t^2\big)} + \frac{a}{c},\ \frac{t}{|c|^2\big(|z + d/c|^2 + t^2\big)} \right) && \text{when } c \neq 0. \\ (z, t) &\mapsto \left(\frac{a}{d}(z + b/a),\ \left|\frac{a}{d}\right| t \right) && \text{when } c = 0. \end{aligned}$$

1.2 Hyperbolic geometry

In the euclidean plane, there is exactly one line through a given point and not meeting a given line disjoint from the point; this is the famous fifth postulate of Euclid. It gradually became clear in the nineteenth century that one can have a self-consistent and interesting geometry where this postulate is not valid — where "parallel" lines are not unique and indeed exist in uncountable abundance. This became known as *hyperbolic geometry*. Though the name was bestowed in connection with conics and projective geometry [Klein 1871, p. 72], it is a doubly felicitous choice, because the Greeks had named the hyperbola after the word for excess (compare "hyperbole", from the same Greek word). Hyperbolic geometric certainly has an excess of lines — and of "room" — compared to euclidean geometry!

Here are some of the salient features that distinguish hyperbolic geometry from the familiar euclidean and spherical geometry.

(i) The angle sum Σ of a hyperbolic triangle Δ satisfies $0 < \Sigma < \pi$; in fact, Σ equals $\pi - \text{area}\,\Delta$. The limiting case $\Sigma = 0$ is achieved by *ideal triangles* whose vertices are "at infinity": we will have more to say about such *ideal vertices* soon (page 14). At the other extreme, the case $\Sigma = \pi$ is the limiting case of hyperbolic triangles of very small area. Indeed, on the infinitesimal scale, hyperbolic geometry is euclidean.

(ii) There are no similarities in hyperbolic space — one cannot scale a figure up or down without changing its angles and shape. It follows, for instance, that all hyperbolic triangles with the same angles are isometric (hyperbolic triangles are "rigid"), and also that the choice of a unit of length is not arbitrary, as in euclidean space; one can privilege a unit having some special property, say the side length of an equilateral triangle whose vertex angles are $\pi/4$.

(iii) For any $0 \le \theta < \pi/(n-2)$ there is a regular n-sided hyperbolic polygon with vertex angles θ. More generally, a necessary and sufficient condition for the existence of an n-sided convex polygon with vertex angles θ_i (with $0 \le \theta_i < \pi$) in clockwise order is that $\sum \theta_i < (n-2)\pi$. The polygon is uniquely determined up to isometry and its area is $(n-2)\pi - \sum \theta_i$.

(iv) Two convex hyperbolic polyhedra that are combinatorially the same with the same dihedral angles and valence 3 at all vertices are isometric [Rivin 1996; Bobenko and Springborn 2004].

(v) The hyperbolic volume V of a ball and the surface area S of its bounding sphere grow exponentially with the hyperbolic radius ρ. The ratio of the surface area to the volume approaches 2 as $\rho \to \infty$.

In short, in the hyperbolic plane and space there are more geometric shapes, they have a tendency toward rigidity, and there is a lot more space in which to build them — in the estimate of Dick Canary, a baseball game played in the hyperbolic plane would require more than 10^{100} ballplayers to provide the same level of outfield coverage as in euclidean space!

Most 2-dimensional abstract surfaces and 3-dimensional manifolds can be modeled using hyperbolic geometry, but not euclidean or spherical geometry. Hyperbolic space is a good place to embed exponentially growing graphs, like a graph representing interconnected web sites. In fact PARC has patented an algorithm for laying out such graphs in $\mathbb{H}^2$ [Lamping et al. 1995]. A different, unpatented, algorithm for laying out graphs in $\mathbb{H}^3$ is presented in [Munzner 1997]. The change of focus from one site to another is effected by a hyperbolic isometry.

By studying the ancient microwave radiation that pervades the universe, astrophysicists hope to get clues about the topology and large-scale curvature of our cosmic

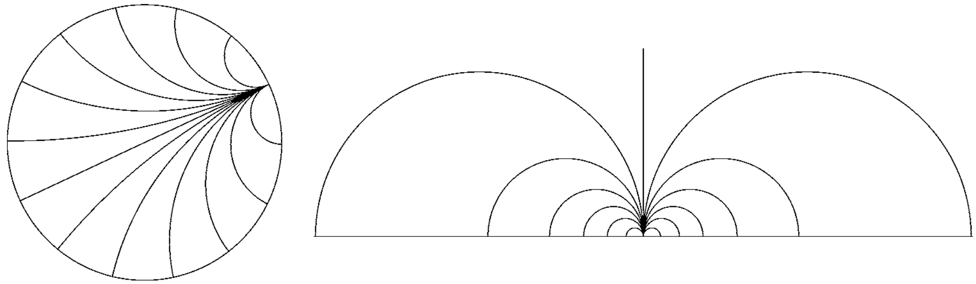


Fig. 1.3. Disk and upper half-plane models of $\mathbb{H}^2$ showing the same geodesics.

home. An earlier proposal that we live in a hyperbolic universe appears to be incompatible with recent data from the Wilkinson Microwave Anisotropy Probe (WMAP), which found the total density (matter plus vacuum energy) to have essentially the value expected for flat space. To the extent that there may be deviation, it is toward a spherical universe (positive curvature); see the discussion in [Weeks 2004]. If the universe is a closed manifold with positive curvature, it can have one of only a few topological types. * To establish that the universe is not simply connected would be astounding!

We now discuss the most commonly used *models* of the hyperbolic plane and of hyperbolic space. These are subsets of $\mathbb{R}^n$ with appropriate riemannian metrics.

The hyperbolic plane

The *upper half-plane model* is $\{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ with the metric

$$ds = \frac{|dz|}{\operatorname{Im} z}.$$

Here $\operatorname{Im} z$ is the notation for the imaginary part. The *unit disk model* is $\{z \in \mathbb{C} : |z| < 1\}$ with the metric

$$ds = \frac{2\,|dz|}{1 - |z|^2}.$$

The two models are equivalent under any Möbius transformation that maps the upper half-plane onto the unit disk. We will denote either one of these models by $\mathbb{H}^2$, the notation for the *hyperbolic plane*. These models have the following properties.

(i) The metrics are *infinitesimally euclidean*; at each point they equal a rescaled euclidean metric. Thus the angle between two curves in the disk or upper half-plane is the same whether measured in the hyperbolic or the euclidean geometry;

* For example, it might conceivably be Poincaré dodecahedral space, the famous first example found by Henri Poincaré of a closed manifold with zero homology which is not homeomorphic to $\mathbb{S}^3$. He had initially believed that such a manifold must be $\mathbb{S}^3$; the example led him to the Poincaré Conjecture. A good explanation of this space and of the classification of spherical three-manifolds can be found in [Thurston 1997].

as a result these models are often called *conformal*. (For other models see Exercise 1-25 and following.)

(ii) $\mathbb{H}^2$ is *complete* in its metric. Every arc tending to the boundary has infinite length.

(iii) The metrics are invariant under any Möbius transformation that maps the model onto itself. In fact these transformations comprise the full group of orientation preserving isometries of the model.

(iv) The hyperbolic lines (geodesics) in the upper half-plane model are semicircles orthogonal to $\mathbb{R}$ and vertical half lines. In the disk model they are diameters and circular arcs orthogonal to $\{|z| = 1\}$.

Hyperbolic space

The *upper half-space model* is $\{(z,t) : z \in \mathbb{C},\ t > 0\}$ with the metric

$$ds = \frac{|d\vec{x}|}{t}, \qquad |d\vec{x}|^2 = |dz|^2 + dt^2.$$

The *ball model* is $\{\vec{x} \in \mathbb{R}^3 : |\vec{x}| < 1\}$ with the metric

$$ds = \frac{2\,|d\vec{x}|}{1 - |\vec{x}|^2}.$$

The two models are equivalent by a Möbius transformation that maps one to the other. Stereographic projection extends to such a Möbius transformation (Exercise 1-11). We will refer to either of these models with its metric as *hyperbolic space* and denote it by $\mathbb{H}^3$.

We repeat our list of properties:

(i) The metrics are infinitesimally euclidean and correctly represent the angles in $\mathbb{H}^3$.

(ii) $\mathbb{H}^3$ is complete in its metric.

(iii) The metrics are invariant under any Möbius transformation that maps the model onto itself. These transformations form the full group of orientation preserving isometries of the models.

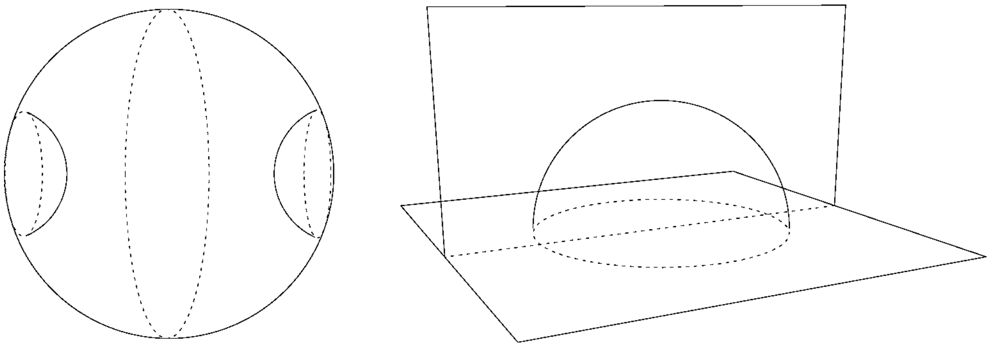


Fig. 1.4. Ball and upper half-space model of $\mathbb{H}^3$ showing geodesic planes.

(iv) The hyperbolic planes in the upper half-space model are hemispheres orthogonal to $\mathbb{C}$ and vertical euclidean half-planes. The lines (geodesics) are semicircles orthogonal to $\mathbb{C}$ and vertical euclidean half-lines. In the ball model the hyperbolic planes are spherical caps orthogonal to the unit sphere, and equatorial planes. The lines are circular arcs orthogonal to the unit sphere, and euclidean diameters.

Restricting the hyperbolic metric to a hyperbolic plane in the model yields the 2-dimensional hyperbolic metric on that plane. Particular cases are the vertical half-plane rising from $\mathbb{R}$ in the upper half-space model and the equatorial plane in the ball model, where the restriction of the metrics give rise to our models of $\mathbb{H}^2$.

Proof of property (iii). For the proof that the Möbius transformations are orientation preserving isometries of the models, see Exercises 1-9 and 1-12. Here we show that there are no other such isometries, concentrating on the hyperbolic plane.

Given three positive distances d_1, d_2, d_3 satisfying the triangle inequality, and a point z on an oriented line $\ell \in \mathbb{H}^2$, there are exactly two triangles with a vertex at z, a side of length d_1 lying on the positive side of ℓ, a side of length d_2 sharing the vertex z, and a third side of length d_3. They are reflections of each other in ℓ and one of the two is uniquely determined if an ordering of the vertices is given and required to give the positive orientation of the triangle they bound.

Given an orientation preserving isometry T, the T-images of three points not on a line are not on a hyperbolic line either. There is a Möbius transformation A such that $A \circ T$ fixes the three points. It then pointwise fixes the sides of the triangle they determine, and then fixes the whole triangle Δ. That is, $T(z) = A^{-1}(z)$, for $z \in \Delta$. If Δ' is a triangle sharing an edge with Δ, there is Möbius transformation A_1 such that $T(z) = A_1^{-1}(z)$ on Δ'. Necessarily $A_1 = A$. Continuing on, building up the whole plane $\mathbb{H}^2$ by adding in succession adjacent triangles, we conclude that $T \equiv A$. □

Proof of property (iv). In view of (iii) we need only prove that the vertical axis ℓ is itself a geodesic. We will work in the upper half-space model. Let ℓ denote the vertical axis rising from $z = 0$. Given $\vec{x} = (z, t) \in \mathbb{H}^3$, define the map $r : \mathbb{H}^3 \to \ell$ as $r(\vec{x}) = (0, t)$. This map is called a *retraction* since in the hyperbolic distance $d(r(\vec{x}), r(\vec{y})) \le d(\vec{x}, \vec{y})$. There is equality if and only if both $\vec{x}, \vec{y}$ lie on a vertical line. This is an immediate consequence of the differential inequality

$$ds^2 = \frac{dx^2 + dy^2 + dt^2}{t^2} \ge \frac{dt^2}{t^2}.$$

Now suppose $\gamma(u)$, with $0 \le u \le 1$, is a differentiable path both of whose endpoints lie on ℓ. Its length strictly exceeds the length of $r(\gamma)$, unless the path is the segment on ℓ between its endpoints. That is, ℓ is a geodesic: the unique shortest path between two points lying on ℓ is the segment of ℓ between the two points. Therefore all images of ℓ by the isometries are also geodesics. In particular, through any two points there passes a unique geodesic.

Likewise the vertical half-plane resting on $\mathbb{R}$ is a hyperbolic plane: the geodesic through any two points of the plane also lies in the plane. Therefore the totality of images under the isometries are the totality of hyperbolic planes. Any three distinct points, not on a line, uniquely determine a hyperbolic plane through them. □

Euclidean circles in $\mathbb{H}^2$ and euclidean circles and spheres in $\mathbb{H}^3$ are also hyperbolic circles and spheres. This is seen by starting with circles and spheres in the disk and ball models which are centered at the origin. The image of a circle or sphere under any Möbius transformation is again a euclidean circle or sphere, if no point on it gets sent to ∞. Conversely any circle or sphere can be sent by a Möbius transformation to one centered at the origin.

However the hyperbolic center is *not* the euclidean center, except for the circles and spheres with center at the origin in the disk and ball models (Exercise 1-4).

1.3 The circle or sphere at infinity

From the point of view of the hyperbolic metric, the models have no boundary: the metric is complete, and hyperbolic straight lines extend forever, though equal hyperbolic distances are represented by increasingly smaller euclidean distances in the model as one approaches the edge (the unit circle in the disk model, etc.)

However, it is useful to regard the edge of the model as a sort of "conformal boundary" in a way that will be explained shortly. This boundary is denoted by $\partial\mathbb{H}^2$ ($=\mathbb{S}^1$ or $\mathbb{R}\cup\{\infty\}$ for the hyperbolic plane) and by $\partial\mathbb{H}^3$ ($=\mathbb{S}^2$ or $\mathbb{C}\cup\{\infty\}$) for hyperbolic space. Another common designation is $\mathbb{S}_\infty$, for the *circle or sphere at infinity*. If we fix a point in $\mathbb{H}^3$, we can also identify $\partial\mathbb{H}^3$ with the *visual sphere* of rays emanating from this point.

In $\mathbb{H}^2$ or $\mathbb{H}^3$, each hyperbolic line determines two "endpoints" on the boundary. Conversely, two distinct boundary points uniquely determine a line. Distinct lines may share an endpoint — indeed, a way to define the sphere at infinity *intrinsically*, without reference to a model, is by taking all oriented geodesics (parametrized by arclength) and defining as equivalent any two that remain within a bounded distance of each other as $t\to\infty$; the set of equivalence classes is $\mathbb{S}_\infty$.

Two hyperbolic lines intersect in at most one point. In $\mathbb{H}^2$, they intersect if and only if their endpoints alternate on $\partial\mathbb{H}^2$. Given a line ℓ and a point $z\notin\ell$ in $\mathbb{H}^2$, there are infinitely many lines through z which do not meet ℓ — unlike the case of the euclidean plane! These are the "parallel lines" of the hyperbolic plane. Among all these parallel lines, there are two that share an endpoint with ℓ.

In $\mathbb{H}^3$ each hyperbolic plane P is bounded by a circle on $\partial\mathbb{H}^3$ (which may be realized as a euclidean line on the boundary of the upper half-space model). Conversely each circle on $\mathbb{S}^2=\partial\mathbb{H}^3$ determines one such plane.

The isometries of $\mathbb{H}^3$ extend to $\partial\mathbb{H}^3$ as conformal automorphisms, that is, as Möbius or anti-Möbius transformations (depending on whether the isometry preserves or reverses orientation).

As mentioned earlier, the set of geodesic rays from a given point $\vec{x}\in\mathbb{H}^3$ can be identified with $\partial\mathbb{H}^3$. Any hyperbolic plane not through $\vec{x}$ subtends a *solid angle* at $\vec{x}$

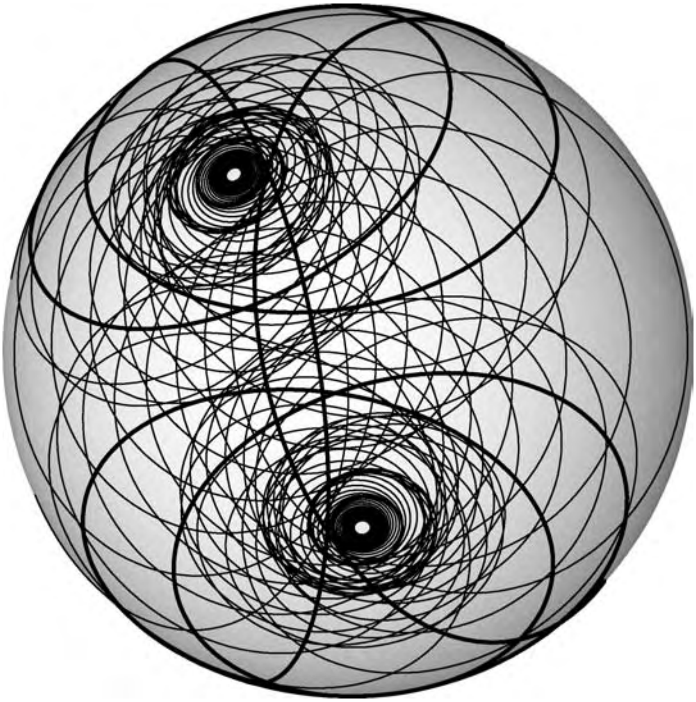


Fig. 1.5. Outer circles: Isometric circles (see page 19) on $\mathbb{S}^2$ of numerous elements of a cyclic group generated by a loxodromic with approximate trace $1.92 + .03i$.

of $< 2\pi$. That is, on a tiny sphere of radius ϵ about $\vec{x}$, the intersection of the sphere with the rays from $\vec{x}$ to the plane fill out a surface area strictly less than $2\pi\epsilon^2$, less than half the area of the sphere. In contrast, in euclidean space any plane subtends exactly a solid angle 2π.

If we lived in hyperbolic space, what we would see as flat lines and planes would automatically be the hyperbolic geodesics, since light would travel along hyperbolic geodesics. If we stood on a plane P, we would see the "circle at infinity" that supports a plane P as the *horizon* of P.

In practice, we have to view hyperbolic space from the outside, from euclidean space using one of our models. We then see the euclidean lines and planes as flat while most of the hyperbolic ones look curved. Looking at the disk or ball model from the outside, we also see the entire circle or sphere at infinity.

From the outside, $\partial\mathbb{H}^3$ is full of circles corresponding to elements of discrete groups of isometries, the *outer circles* of the book title. (See Figure 1.5.) The action of isometries on geodesic planes in $\mathbb{H}^3$ is paired with the corresponding action on the outer circles.

An elliptic transformation T has an *axis of rotation* inside $\mathbb{H}^3$. It is the hyperbolic line connecting its fixed points on $\mathbb{S}^2$. The axis is pointwise fixed by T.

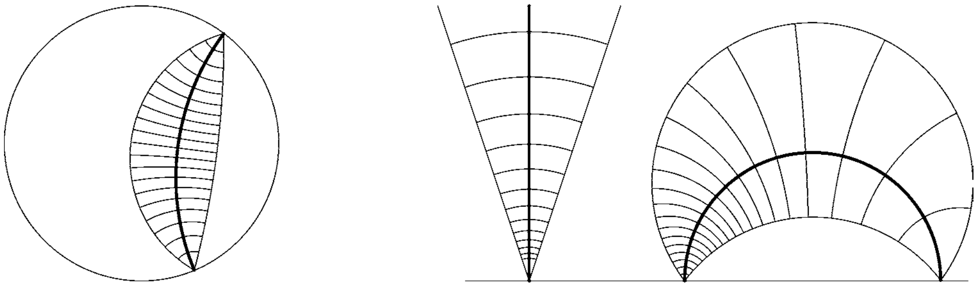
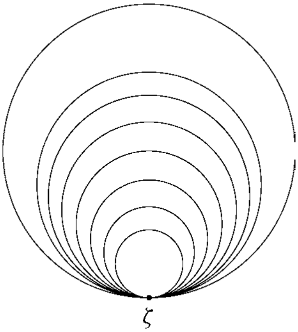


Fig. 1.6. Invariant tubes viewed in cross section, in the ball model and the upper half-space. The axis of the transformation is shown thicker in each case. The transversal lines represent discs orthogonal to the axis; all these disks (within the same tube) are congruent.

A loxodromic transformation T likewise has an *axis* in $\mathbb{H}^3$. It too is the hyperbolic line connecting the fixed points. T maps the line onto itself, moving each point toward the attracting fixed point. If T is in standard form, $\lambda^2 z$, with $|\lambda| > 1$, the axis is the vertical half-line $z = 0$ in upper half-space. The hyperbolic distance between any pair of points z, $T(z)$ on the axis is $d = 2\log|\lambda|$, or $2\cosh(d/2) = |\lambda| + |\lambda^{-1}| \geq |\tau_T|$.

Both elliptic and loxodromic transformations in $\mathbb{H}^3$ leave invariant not just the axis, but also each of a family of surfaces equidistant from the axis. These surfaces are particularly easy to visualize in upper half-space when the transformation is in standard form: the surface is a euclidean cone with vertex at $(z, t) = (0, 0)$ and a vertical axis (the half-line $z = 0$). When both endpoints of the axis line on the plane $t = 0$, the euclidean shape of the surfaces is a tube, tapering to a cone at each endpoint (Figure 1.6). Note that though we often describe features of hyperbolic space by talking about their euclidean shapes in the model — and this mixture is almost inevitable — you should strive to visualize each object both intrinsically (the tube has constant diameter) and in terms of the model (the tube looks like a cone or a crescent).

A parabolic transformation P has no axis, since there is only one fixed point. P does have invariant surfaces, each mapped to itself (in the spirit of the tubes of the previous paragraph); they look like euclidean spheres, and are called *horospheres*. (Watch out: horospheres are *not* hyperbolic spheres! See Exercise 1-33.) All horospheres of a parabolic transformation P are tangent to one another and to the sphere at infinity at the fixed point ζ of P. The region of $\mathbb{H}^3$ cut off by a horosphere is called a *horoball*. In the upper half-space model, there is an exceptional case, when the fixed point ζ is at infinity (say for $P(z) = z + 1$): then the horospheres are euclidean planes $\{(z, t) : t = \text{constant}\}$, and the horoballs are the half-spaces above these planes. For parabolic transformations of $\mathbb{H}^2$ the corresponding objects in are called *horocycles* and *horodisks*.



At a parabolic fixed point $\zeta \in \mathbb{S}^2$, there is a double family of mutually tangent circles at ζ that bound disjoint open disks and are invariant under the parabolic transformation. In the case $\zeta = \infty$ and $P(z) = z + a$, with $a \in \mathbb{R}$, they are the family $\{z : \operatorname{Im} z = \pm s,\ s \neq 0\}$. More generally if the flow associated with P has the vector direction $a \in \mathbb{C}$, the pairs of horocycles are $\{z : \operatorname{Im} \bar{a}z = \pm s,\ s \neq 0\}$.

The prefix "horo" comes form the Greek word for "limit". Fix a point $O \in \mathbb{H}^3$. Take the *hyperbolic* sphere σ_x centered at $x \in \mathbb{H}^3$ and passing through O. As $x \to \zeta \in \partial\mathbb{H}^3$, the limit of σ_x is the horosphere at ζ passing through O.

We now take up the study of triangles. As already mentioned the area of a triangle is equal to the "angle deficit" $\pi - \sum \theta_i$, where the θ_i are the vertex angles; see Exercise 1-6 for a proof. Thus the greatest area a triangle can have is π, which happens when all vertices have "angle zero" — this is really a limiting case, when the vertices are no longer points in hyperbolic space by in the sphere at infinity. A point in the sphere at infinity is also called an *ideal point*, and so triangles whose vertices are at infinity are *ideal triangles*. Given two ideal triangles and a labeling of the respective vertices in the positive direction, there is a unique isometry that takes one to the other, matching the designated labeling.

Theorem 1.3.1 (All triangles are thin). *Any point ξ on a side of a hyperbolic triangle Δ is within distance* $\log(1 + \sqrt{2}) = \operatorname{arcsinh} 1$ *from one of the two other sides. The distance attains its maximum only for an ideal triangle, with ξ of equal distance from the two other sides.*

Any point inside a hyperbolic triangle is within distance $\log(1 + \sqrt{2}) = \operatorname{arcsinh} 1$ *of one of the sides.*

Proof. We work in the upper half-plane model. We may assume by changing the position of Δ in $\mathbb{H}^2$ by an orientation preserving isometry that the side $[p, q]$ of $\Delta = (p, q, r)$ containing ξ lies on the unit semicircle centered at the origin, Δ lies above this semicircle, and the side $[p, r]$ lies on the vertical euclidean line through

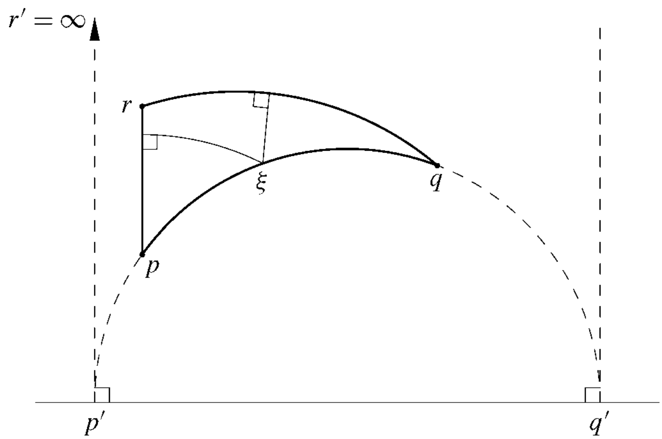


Fig. 1.7. Universal thinness of triangles.

p. Here take p to be the left vertex on the semicircle. Assume at the start that none of the vertices is ideal. (See Figure 1.7.)

We start by showing that we may replace r by $r' = \infty$. As r goes up along the vertical line from p, the distance of ξ to $[p, r]$ increases or (once r is no longer the closest point to ξ) remains the same. The distance of ξ to the side $[q, r']$, too, either equals the distance $q\xi$ or the length of the perpendicular from ξ to $[q, r']$; therefore it exceeds the distance from ξ to $[q, r]$, since $[q, r]$ separates $[q, r']$ from ξ in the new triangle (p, q, r'). Thus the minimum distance of ξ to the sides increases when we more r to $r' = \infty$, so that Δ becomes a triangle with an ideal vertex.

Next consider what happens as q slides down the semicircle to its right endpoint $q' \in \mathbb{R}$. The distance of ξ to $[q, \infty]$ strictly increases as q changes to q'; in the new triangle (p, q', ∞) the side $[q, \infty]$ separates $[q', \infty]$ from ξ. Similarly the distance of ξ to $[p, \infty]$ strictly increases as p slides down the semicircle to its left endpoint $p' \in \mathbb{R}$.

So now we have an ideal triangle with ξ on the side that is now the full semicircle. The minimal distance of ξ to the two vertical sides is greatest when ξ is the symmetric point $\xi = i$. Finally we have to compute the distance from ξ to one of the vertical sides. There is exactly one semicircle C' through $\xi = i$ with center at $z = p'$ which is orthogonal to the vertical line $[p', \infty]$. In polar coordinates at p', the orthogonal segment is the arc $0 \leq \theta \leq \pi/4$ of C', if θ is measured from the vertical. The length of this segment is

$$\int_0^{\pi/4} \frac{\rho\, d\theta}{\rho \cos\theta} = \log(\sqrt{2} + 1).$$

Here the radius ρ of C' doesn't enter — the map $z \mapsto kz$, $k > 0$ (a euclidean similarity) is a hyperbolic isometry.

Given a point z in a triangle Δ and a pair of sides, divide Δ by a geodesic arc between the designated sides and passing through z. Application of what we just proved shows that z is within distance arcsinh 1 of the two designated sides. Then repeat the argument with an arc through z from the third side. □

A related fact is described in Exercise 1-17.

1.4 Gaussian curvature

The hyperbolic plane is a simply connected surface with a complete riemannian metric of constant negative gaussian curvature. It is usually taken (by multiplying the metric by the appropriate constant) to be -1, as we have done in §1.2. The purpose of this section is to explain the meaning of the expression "gaussian curvature -1".

Using the disk model of $\mathbb{H}^2$ and polar coordinates (r, θ) based at the origin, we begin with the following computations: the hyperbolic radius ρ of the circle of euclidean radius $R < 1$ centered at the origin, its hyperbolic area A, and its hyperbolic circumference C. Our results are as follows:

$$\rho = \int_0^R \frac{2\,dr}{1-r^2} = \log\frac{1+R}{1-R}, \qquad R = \tanh\frac{\rho}{2} = \frac{\cosh\rho - 1}{\sinh\rho},$$

$$A(\rho) = \int_{\theta=0}^{2\pi}\int_{r=0}^{R} \frac{4r\,dr\,d\theta}{(1-r^2)^2} = \frac{4\pi R^2}{1-R^2} = 2\pi(\cosh\rho - 1),$$

$$C(\rho) = \int_0^{2\pi} \frac{2R\,d\theta}{1-R^2} = \frac{4\pi R}{1-R^2} = 2\pi\sinh\rho.$$

$$C(\rho)^2 = A(\rho)^2 + 4\pi A(\rho).$$

Thus both the area and the circumference grow exponentially with the hyperbolic radius; more than 63% of the surface of any hyperbolic disk is within 1 unit of the boundary; the ratio $A(\rho)/C(\rho)$ equals $R < 1$ and in particular approaches 1 as $\rho \to \infty$.

(For the record, the analogous formulas for the volume $V(\rho)$ of a hyperbolic ball of hyperbolic radius ρ and its surface area $S(\rho)$ are:

$$V(\rho) = 8\int_0^{2\pi} d\theta \int_0^{\pi} \sin\phi\,d\phi \int_0^R \frac{r^2\,dr}{(1-r^2)^3} = \pi(\sinh 2\rho - 2\rho),$$

$$S(\rho) = 4\int_0^{2\pi}\int_0^{\pi} \frac{R^2\sin\phi\,d\phi\,d\theta}{(1-R^2)^2} = 4\pi\sinh^2\rho.$$

Thus $2V(\rho) < S(\rho)$ and $\lim_{\rho\to\infty} 2V(\rho)/S(\rho) = 1$.)

Now consider a smooth riemannian surface, a point z on the surface, and, for $\rho > 0$ variable, the disk of radius ρ around that point (in the given metric). The gaussian curvature K_0 at z can be characterized by the following properties involving the limiting behavior of the area $A(\rho)$ and circumference $C(\rho)$ of such disks, compared with their euclidean counterparts [Struik 1950, §4.3]:

$$K_0 = -3\frac{d^2}{d\rho^2}\left(\frac{C(\rho)}{2\pi\rho}\right)_{\rho=0} = \frac{3}{\pi}\lim_{\rho\to 0}\frac{2\pi\rho - C(\rho)}{\rho^3},$$

$$K_0 = -6\frac{d^2}{d\rho^2}\left(\frac{A(\rho)}{\pi\rho^2}\right)_{\rho=0} = \frac{12}{\pi}\lim_{\rho\to 0}\frac{\pi\rho^2 - A(\rho)}{\rho^4}.$$

In particular negative curvature is characterized by the property that $C(\rho) > 2\pi\rho$ for all small values of R. Or by the property that $A(\rho) > \pi\rho^2$. This is confirmed for $\mathbb{H}^2$ from the formulas for area and circumference above. Contrast this with the corresponding properties of euclidean space.

Here is a construction of a surface with discrete negative curvature: Take equilateral (euclidean) triangles with unit side lengths. Of course these tessellate the euclidean plane; six are arranged about each vertex. Instead form a polyhedral surface by placing seven triangles about each vertex. This forms a polyhedral surface which is flat except at the vertices. In a polyhedral surface, each vertex v has a discrete curvature defined by $2\pi - \sum \theta_i$, where the θ_i are the vertex angles at v of the triangles sharing v. In this case the curvature at each vertex is $-\pi/3$. On our surface, the "circle" of radius $R = 1$ about a vertex has circumference 7, larger than the euclidean circumference of 6 when there are six triangles about each vertex.

If $u\,|dz|$, with $u(z) > 0$, is a conformal metric on a Riemann surface, the gaussian curvature of the metric can be defined in terms of the laplacian as

$$K_u = -\frac{\Delta \log u}{u^2} = -\frac{4}{u^2}\frac{\partial^2 \log u}{\partial z\,\partial \bar{z}}.$$

Ahlfors had great success early in his career [1973] with applications involving singular conformal metrics of this form.

If instead we want a model of $\mathbb{H}^2$ with gaussian curvature $-c < 0$, take the disk $|z| < 1/\sqrt{c}$ with the metric

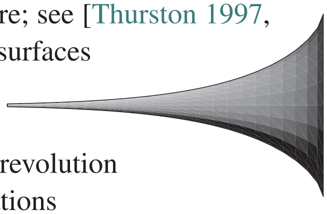
$$\frac{2\,|dz|}{1 - c\,|z|^2}.$$

Gauss originally defined the curvature as follows. Suppose $S \subset \mathbb{R}^3$ is an embedded surface and $p \in S$. Draw a simple closed curve $c \subset S$ enclosing a region $D \subset S$ containing p. Interpret each exterior unit normal vector $\vec{N}$ (determined by the right-hand rule) at a point of D as a vector from $(0, 0, 0)$ to the unit 2-sphere $\mathbb{S}^2$. As $\vec{N}$ ranges over all possibilities, a certain region $\Omega \subset \mathbb{S}^2$ is filled out. Gauss defined the *total curvature* of D to have absolute value $A(\Omega)$, the area of Ω. The sign is determined as follows. As $\vec{N}$ runs over c in the positive direction (D to its left), use $+$ if the corresponding $\vec{N}$ runs over $\partial\Omega$ also in the positive direction (Ω to its left); otherwise use $-$. Thus the total curvature of a region in a plane is zero, while the total curvature of a hemisphere is $2\pi\,\sin(\pi/2)$. Gauss defined the *curvature* of S at the point p as

$$\lim_{D \searrow \{p\}} \frac{\pm A(\Omega)}{A(D)}.$$

Gaussian curvature is an intrinsic property of a surface — although the definition just given is for surfaces embedded in $\mathbb{R}^3$, Gauss's famous *Theorema Egregium* is that isometric surfaces have the same gaussian curvature at corresponding points. So we can define the curvature for a metric defined on an abstract surface.

Hilbert proved that there is no C^2 surface in $\mathbb{R}^3$ whose metric induced from $\mathbb{R}^3$ is a *complete* metric of constant negative gaussian curvature; see [Thurston 1997, §2.1], for example. There do, of course, exist smooth surfaces embedded in $\mathbb{R}^3$ with constant negative curvature, but they cannot be extended to a complete surface. The most famous example is the *pseudosphere*, a surface of revolution about the x-axis in $\mathbb{R}^3$ described by the parametric equations



$$(u - \tanh u,\ \operatorname{sech} u \cos v,\ \operatorname{sech} u \sin v) \quad \text{for } u \geq 0,\ v \in \mathbb{R}.$$

If we take the subset $\{z : \operatorname{Im} z \geq 1 > 0\}$ of $\mathbb{H}^2$ and quotient it by the cyclic group of hyperbolic isometries generated by $z \to z + 2\pi$, we get a riemannian surface isometric to the pseudosphere (with its inherited metric from $\mathbb{R}^3$); see [Coxeter 1961, p. 378], for example. Conformally, this is a once punctured disk.

The *Gauss–Bonnet formula* for a *simply connected* surface element S of gaussian curvature K, bounded by the union of n smooth arcs meeting with interior angles θ_i at the vertices, is

$$\iint_S K\, dS + \int_{\partial S} \kappa_g\, ds = 2\pi - \sum_{i=1}^{n} (\pi - \theta_i) = \pi(2-n) + \sum_{i=1}^{n} \theta_i, \tag{1.3}$$

where κ_g is the geodesic curvature of the arcs [Struik 1950, §4.8]. For a geodesic arc, the geodesic curvature κ_g vanishes; thus, for example, if $S \subset \mathbb{H}^2$ is a hyperbolic triangle Δ, the formula becomes

$$-\text{area}\, \Delta = -\pi + \theta_1 + \theta_2 + \theta_3.$$

The Gauss–Bonnet formula, and indeed the area formula for triangles directly, can be verified by using Green's formula (see also Exercise 1-6). By breaking more general surfaces into simply connected regions one can apply the formula further.

See Exercise 1-33 for computations of the hyperbolic curvature of horocycles, equidistant arcs to geodesics, and circles.

For hyperbolic 3-space $\mathbb{H}^3$ (or n-space more generally), the normalized metric is characterized by having *sectional curvature* -1: all 2-dimensional planes through a given point have gaussian curvature -1 in the metric induced from that of $\mathbb{H}^3$.

1.5 Further properties of Möbius transformations

The following facts must be part of our repository of basic knowledge.

Commutativity

Lemma 1.5.1. *Let A, B be Möbius transformations* $\neq \text{id}$.

(*i*) *A and B share a fixed point if and only if*

$$\operatorname{tr}(ABA^{-1}B^{-1}) = +2.$$

(*ii*) *Assume that A and B do not share a fixed point. Then $ABA^{-1}B^{-1}$ is parabolic if and only if*

$$\operatorname{tr}(ABA^{-1}B^{-1}) = -2.$$

Proof. The second statement follows directly from the characterization on page 3. The proof of the first is not hard is one takes one of the transformations to be in standard form. □

Lemma 1.5.2. *Let A and B be Möbius transformations distinct from* $\pm\,\text{id}$. *There is equivalence between*:

(i) *A and B commute.*
(ii) *Either A and B have the same set of fixed points, or A and B have order two and each interchanges the fixed points of the other.*
(iii) *Either A and B are parabolic with the same fixed point, or the axes of A and B coincide, or A and B have order two and their axes intersect orthogonally in $\mathbb{H}^3$.*

Again, this can be checked by assuming A to be in standard form, and observing that the general elliptic transformation of order two exchanging 0 with ∞ is $z \mapsto a^2/z$.

Lemma 1.5.3. *Let $k \geq 1$ and let A and B be Möbius transformations, with $A^k \neq \mathrm{id}$. Then $A^k = B^k$ if and only if either $A = B$, or $A = EB$ for an elliptic E whose order divides k and whose set of fixed points is the same as that of B.*

Proof. Sufficiency is obvious, since E and B commute by Lemma 1.5.2.

To prove necessity, we can assume that k is minimal with the property that $A^k = B^k$. Note that taking powers preserves both type and fixed points, except that an elliptic can become the identity. Thus, if A is parabolic, so is A^k, and hence so is B; but the parabolic elements fixing a given point of S_∞ form a torsion-free abelian group (see again Lemma 1.5.2), so $A = B$ in this case. If instead A fixes two points, the same argument (allied to the fact that $A^k \neq \mathrm{id}$) again shows that A and B commute; hence $(AB^{-1})^k = \mathrm{id}$, so $E = AB^{-1}$ is elliptic and shares the fixed points of A, B. □

Isometric circles and planes

Consider a Möbius transformation A on $\mathbb{C} \cup \infty$. If A does not fix ∞, it has the form

$$A(z) = \frac{az+b}{cz+d}, \quad ad - bc = 1, \ c \neq 0.$$

One may ask, at what points does A preserve the size of (euclidean) tangent vectors, as well as angles? Since $A'(z) = 1/(cz+d)^2$, the set of such points, denoted by

$$\mathcal{I}(A) = \{z \in \mathbb{C} : |A'(z)| = 1\} = \{z \in \mathbb{C} : |cz+d| = 1\},$$

is a circle. We call it the *isometric circle* of A. Its center and radius satisfy

$$\operatorname{center} \mathcal{I}(A) = -\frac{d}{c} = A^{-1}(\infty), \quad \operatorname{radius} \mathcal{I}(A) = \frac{1}{|c|}.$$

Because A maps circles to circles, the restriction of A to $\mathcal{I}(A)$ is a euclidean isometry onto the circle $\mathcal{I}(A^{-1})$ of the same radius. Also, $|A'(z)| > 1$ for z in the interior of $\mathcal{I}(A)$, and $|A'(z)| < 1$ for z in the exterior.

Now consider the same transformation A, regarded as an isometry of $\mathbb{H}^3$ in the upper half-space model (recall that $\mathbb{C} \cup \infty = \partial\mathbb{H}^3$). The *isometric plane* of A is likewise defined as the set of points where the jacobian preserves length:

$$\{\vec{x} : |A'(\vec{x})| = 1\}. \tag{1.4}$$

Explicitly, it is the hemisphere

$$\{\vec{x} = (z,\ t),\ t > 0 : |cz + d|^2 + |c|^2 t^2 = 1\}.$$

rising from the isometric circle $\mathcal{I}(A) \subset \partial\mathbb{H}^3$. When the context is clear, we will use the notation $\mathcal{I}(A)$ interchangeably for the isometric circle and the isometric plane.

There is also an isometric circle and plane for the ball model, with the same defining equation (1.4) as for the upper half-space model. However, the isometric circle in one model does not usually map to the isometric circle in the other under a hyperbolic isometry conjugating the two models; in fact, even within the same model, a conjugating isometry U need not map $\mathcal{I}(A)$ to $\mathcal{I}(UAU^{-1})$, as is the case with the axis and fixed points. Although not intrinsic, the notion of isometric circles and planes is nonetheless useful because of its metric properties. It was introduced by L. R. Ford.

Here is another description of the isometric planes (see Exercises 1-10 and 3-4).

Lemma 1.5.4.

(i) *In the ball model, the isometric plane for A is the perpendicular bisector of the line segment* $[0,\ A^{-1}(0)]$, *where* 0 *denotes the origin of the ball.*

(ii) *In the upper half-space model, if ∞ is not a fixed point of A, the isometric plane results from the following construction. There is exactly one horosphere $\mathcal{H}$ at ∞ such that the horosphere $A^{-1}\mathcal{H}$ at $A^{-1}(\infty)$ is tangent to $\mathcal{H}$. The line ℓ between ∞ and $A^{-1}(\infty)$ goes through the point of tangency and is orthogonal to the two horospheres. The isometric plane is the unique plane through the point of tangency and orthogonal to ℓ.*

We summarize here the properties of isometric planes and circles in the upper half-space model. Refer to Figure 1.8 for examples.

Proposition 1.5.5. *Let A be a Möbius transformation of the upper half-space model, not fixing ∞. Let $\mathcal{B}(A)$ be the closed disk bounded by the isometric circle $\mathcal{I}(A)$ in $\mathbb{C}$, and let $\mathcal{E}(A)$ be the closure of its exterior* (*including* ∞).

(1) *A sends $\mathcal{I}(A)$ to $\mathcal{I}(A^{-1})$, $\mathcal{B}(A)$ to $\mathcal{E}(A^{-1})$, and $\mathcal{E}(A)$ to $\mathcal{B}(A^{-1})$. If $\vec{x} = (z, t)$ lies on the isometric plane $\mathcal{I}(A)$, then $A(\vec{x}) = (A(z), t)$ lies on $\mathcal{I}(A^{-1})$.*

(2) *$\mathcal{I}(A) = \mathcal{I}(A^{-1})$ if and only if $\tau_A = 0$.*

(3) *The intersection of circles $\mathcal{I}(A) \cap \mathcal{I}(A^{-1})$ consists of two points if and only if $0 < |\tau_A| < 2$. If A is elliptic, these intersection points are the fixed points and the corresponding isometric planes intersect in the axis of rotation.*

(4) *$\mathcal{I}(A)$ and $\mathcal{I}(A^{-1})$ are tangent if and only if A is parabolic, in which case the tangency point is the fixed point.*

(5) *$\mathcal{I}(A)$ and $\mathcal{I}(A^{-1})$ are disjoint if and only if $|\tau_A| > 2$.*

(6) *If A is loxodromic, $\mathcal{B}(A) \cap \mathcal{E}(A^{-1})$ contains the repelling fixed point of A, and $\mathcal{B}(A^{-1}) \cap \mathcal{E}(A)$ its attracting fixed point.*

(7) *If U fixes ∞, then $\mathcal{I}(UAU^{-1}) = U(\mathcal{I}(A))$. If U is a euclidean translation, $\mathcal{I}(UA) = \mathcal{I}(A)$ and $\mathcal{I}(AU^{-1}) = U\mathcal{I}(A)$.*

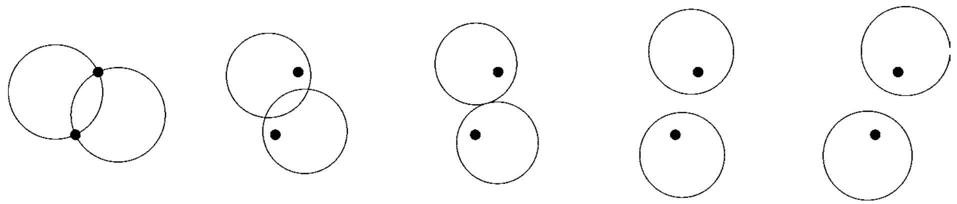


Fig. 1.8. Isometric circles for various transformations A that have the same fixed points. From left to right, the elliptic case ($\tau_A \in (-2,2)$), the loxodromic case with $\tau_A \notin \mathbb{R}$ and respectively $|\tau_A| < 2$, $|\tau_A| = 2$ and $|\tau_A| > 2$, and finally the loxodromic case with τ_A real.

(8) *If A preserves a circle in $\mathbb{S}^2$, $\mathcal{I}(A)$ is orthogonal to that circle.*

Proof. (1) The Jacobian determinants or derivatives are related:

$$\left(BA^{-1}\right)'(Az) = B'(z)/A'(z).$$

(2) A normalized matrix $A = \pm A^{-1}$ if and only if either $A = \pm\,\mathrm{id}$ or $\tau_A = 0$. The isometric circle $|-cz+a| = 1$ of A^{-1} is identical to that of A if and only $a+d=0$.

(3)–(5) The distance between the centers of the isometric circles is

$$\left|\frac{a}{c} + \frac{d}{c}\right| = \frac{|a+d|}{|c|}.$$

Since the radius of the circles is $1/|c|$ they intersect whenever the distance between centers is less than $2/|c|$ and are tangent when there is equality.

Now the distance between the centers is exactly $2/|c|$ when A is parabolic, less than $2/|c|$ when A is elliptic, and can have any positive value when A is loxodromic. Only when the loxodromic satisfies $|\tau_A| > 2$ does the distance between centers exceed $2/|c|$ so that the circles are disjoint.

(6) The derivative $|A'|$ is greater than 1 at the repelling fixed point, and less than 1 at the attracting one, when these points are finite. (In contrast, at a finite elliptic or parabolic fixed point ζ, $|A'(\zeta)| = 1$.)

(7) This is a direct computation, or an application of the chain rule.

(8) If A preserves $\mathbb{R} \cup \infty$, its normalized form has real or purely imaginary entries, the latter case if A interchanges the upper and lower half-planes. Therefore the center of the isometric circle is real, so $\mathcal{I}(A)$ is orthogonal to $\mathbb{R}$. If A maps the unit disk onto itself, it has the following form (Exercise 1-2):

$$A = e^{i\theta}\,\frac{z-a}{1-\bar{a}z}.$$

From this we compute that $\mathcal{I}(A)$ has center $1/\bar{a}$ and squared radius $(1-|a|^2)/|a|^2$. This implies that $\mathcal{I}(A)$ is orthogonal to the unit circle. If A interchanges the two sides of the unit circle, it can be expressed by replacing z by $1/z$ in the formula and proceeding in the same way. The general transformation A is conjugate to one we have considered via a transformation that fixes ∞. $\square$

Trace identities

Here we present the common trace identities that help form the bridge between the algebra of matrices and hyperbolic geometry. See also Exercise 1-20.

Lemma 1.5.6. *Let X and Y be 2×2 complex matrices of determinant one.*

(*i*) $\operatorname{tr}(XY^{-1}) = \operatorname{tr}(X)\operatorname{tr}(Y) - \operatorname{tr}(XY)$.
(*ii*) $\operatorname{tr}(XYX^{-1}Y^{-1}) + 2 = \operatorname{tr}^2(X) + \operatorname{tr}^2(Y) + \operatorname{tr}^2(XY) - \operatorname{tr}(X)\operatorname{tr}(Y)\operatorname{tr}(XY)$.
(*iii*) $\operatorname{tr}(XYX^{-1}Y^{-1}) - 2 = \big(\operatorname{tr}(X) - \operatorname{tr}(Y)\big)^2 - \big(\operatorname{tr}(XY) - 2\big)\big(\operatorname{tr}(XY^{-1}) - 2\big)$.
(*iv*) *If* $\operatorname{tr}^2(X) \neq 4$ *then*

$$\frac{\operatorname{tr}(X^m Y X^{-m} Y^{-1}) - 2}{\operatorname{tr}^2(X^m) - 4} = \frac{\operatorname{tr}(XYX^{-1}Y^{-1}) - 2}{\operatorname{tr}^2(X) - 4}.$$

(*v*) *If* $[X, Y] = XYX^{-1}Y^{-1}$ *is parabolic and* X, Y *do not share a fixed point, so that* $\operatorname{tr}(XYX^{-1}Y^{-1}) = -2$, *then*

$$\operatorname{tr}(X)\operatorname{tr}(Y)\operatorname{tr}(XY) = \operatorname{tr}^2(X) + \operatorname{tr}^2(Y) + \operatorname{tr}^2(XY),$$
$$\operatorname{tr}(XY)\operatorname{tr}(XY^{-1}) = \operatorname{tr}^2(X) + \operatorname{tr}^2(Y).$$

Conversely, either of these two identities implies $\operatorname{tr}[X, Y] = -2$.

Remark 1.5.7. The first equation in (v) is called the *Markov identity*. Markov proved that for the equation $xyz = x^2 + y^2 + z^2$, the only integer solutions (called Markov triples) are provided by the traces of group elements $X, Y, Z = XY$ in the modular group (Exercise 2-9), with $\operatorname{tr}[X, Y] = -2$. If (u, v, w) is a Markov triple, so are $(u, v, uv - w)$, $(u, uw - v, w)$, $(vw - x, v, w)$. A famous unsolved problem in number theory is Markov's conjecture that if (x, y, z), (x', y', z) are Markov triples, with $x \le y \le z$ and $x' \le y' \le z'$, then $x = x'$ and $y = y'$. See [Bowditch 1998] and [Goldman 2003] for more detail.

Proof. To verify (i), (ii) and (iii), apply a conjugacy to convert Y to standard form:

$$Y = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \text{ or } \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}, \qquad X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \ ad - bc = 1.$$

The identities are now easily verified. In particular we find that for Y normalized we have, depending on whether Y is nonparabolic and parabolic,

$$\operatorname{tr}(XYX^{-1}Y^{-1}) = 2 - bc(\lambda - \lambda^{-1})^2 \text{ or } 2 + c^2\lambda^2.$$

Thus the commutator cannot have trace $+2$ unless $b = 0$, $c = 0$, or $X = \pm I$. All three possibilities are excluded by the hypotheses.

The Markov identity in (v) can be regarded as a quadratic equation for $w = \operatorname{tr}(XY)$ in terms of the coefficients $\operatorname{tr}(X)$ and $\operatorname{tr}(Y)$. The two solutions are $w = \operatorname{tr}(XY)$ and $w = \operatorname{tr}(XY^{-1})$. This is the reason for the second identity in (v). If the transformations corresponding to X and Y are loxodromic and preserve the upper half-plane, and their matrix representations are chosen so that the traces are positive, then $\operatorname{tr}(XY)$ will automatically be positive as well.

For item (iv), put X instead of Y in standard form. It is enough to verify the formula for $m = 2$; the general case follows by induction. The ratio has the constant value $-bc$ if $Y = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. $\square$

1.6 Exercises and explorations

1-1. For a Möbius transformation T of $\mathbb{C} \cup \infty$ with normalized matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, prove:

(i) $|T(z) - T(w)| = |z - w|\sqrt{|T'(z)|}\sqrt{|T'(w)|}$.

(ii) T preserves the upper half-plane if and only all of a, b, c, d are real. (Thus the group of orientation preserving isometries of the hyperbolic plane is $\mathrm{PSL}(2, \mathbb{R})$.) If such a T is loxodromic or parabolic, its fixed points lie in $\mathbb{R}$. If T is elliptic, its fixed points are symmetric about $\mathbb{R}$ under reflection, but do not lie in $\mathbb{R}$. Moreover, $\operatorname{Im} T(z) = (\operatorname{Im} z) \cdot |T'(z)|$.

(iii) T preserves the right half-plane if and only if $T = \left(\begin{smallmatrix} a' & ib' \\ -ic' & d' \end{smallmatrix}\right)$ with $a', b', c', d' \in \mathbb{R}$ and $a'd' - b'c' = 1$.

(iv) Find conditions on a, b, c, d for T to preserve the unit disk. (*Hint:* Conjugate by a Möbius transformation taking $-1, 1, \infty$ to $-1, 1, i$.) Prove an alternative characterization: T preserves the unit disk if and only if it can be written as

$$T(z) = e^{i\theta} \frac{z - z_0}{1 - \bar{z}_0 z}, \qquad |z_0| < 1.$$

Moreover, such a T satisfies $|T'(z)|(1 - |z|^2) = (1 - |T(z)|^2)$.

1-2. Two points are said to be symmetric in a circle or straight line if reflection in the circle or line carries one point to the other. Thus z and $\bar{z}$ are symmetric in $\mathbb{R}$, while z and $1/\bar{z}$ are symmetric in the unit circle centered at the origin. Verify that the formula for symmetric points ζ, ζ^* with respect to the circle $\{|z - a| = R\}$ is

$$\zeta^* - a = \frac{R^2}{\bar{\zeta} - \bar{a}}.$$

This map extends to a reflection about the corresponding hyperbolic plane in $\mathbb{H}^3$.

(i) Prove that a Möbius transformation maps points symmetric in a circle/line to points symmetric in the image circle/line. Hence the extension to $\mathbb{H}^3$ preserves symmetry in planes. (See [Ahlfors 1978].)

(ii) If C_1, C_2 are disjoint circles and Δ is the region they bound on $\mathbb{S}^2$, show how to find a Möbius transformation that sends Δ to an annulus centered at $z = 0$ with C_2 sent to the outer circle of radius 1.

(iii) Suppose $D_2 \subset D_1$ are disks centered at $z = 0$ of radii $r_2 < r_1$ respectively. Let T be any Möbius transformation such that $T^{-1}(\infty)$ is not in the closure of D_1. Denote the radii of $T(D_2) \subset T(D_1)$ by r_2', r_1'. Show that

$$\frac{r_1'}{r_2'} \geq \frac{r_1}{r_2}.$$

When is there equality?

1-3. *Liouville measure.* Suppose ℓ is a hyperbolic line in the disk model, with end points a, b. Suppose p, q are two points on ℓ so labeled that p separates a and q. Show that the hyperbolic distance $d(p,q)$ between p and q is given in terms of the cross ratio by

$$d(p,q) = \log(a,b,p,q) = \log \frac{(q-a)(b-p)}{(p-a)(b-q)}.$$

Suppose ℓ_1, ℓ_2 are hyperbolic lines which intersect in $\mathbb{H}^2$ with angle θ. Suppose their end points a_i, b_i, $i = 1, 2$, are arranged clockwise $[a_1, a_2, b_1, b_2]$ around $\partial\mathbb{H}^2$. Prove that

$$(a_1, a_2, b_1, b_2) = \cos^2 \frac{\theta}{2} = \frac{\cos\theta + 1}{2}.$$

The *Liouville measure* is a measure on the space $G(\mathbb{H}^2)$ of unoriented geodesics in $\mathbb{H}^2$. In terms of endpoints on $\mathbb{S}^1 = \partial\mathbb{H}^2$,

$$G(\mathbb{H}^2) = \left(\mathbb{S}^1 \times \mathbb{S}^1 \setminus \{\text{diagonal}\}\right)/\mathbb{Z}_2,$$

which is topologically a Möbius band (Exercise 4-15). The Liouville distance L between nonintersecting geodesics γ_1, γ_2 with endpoints (a, b) and (c, d) is

$$L(\gamma_1, \gamma_2) = \left| \log \left| \frac{(a-c)(b-d)}{(a-d)(b-c)} \right| \right|.$$

For any isometry T, $L(T\gamma_1, T\gamma_2) = L(\gamma_1, \gamma_2)$.

The infinitesimal form of L is

$$\frac{d\alpha\, d\beta}{|e^{i\alpha} - e^{i\beta}|^2} = \frac{d\alpha\, d\beta}{4} \sin^2 \frac{\alpha - \beta}{2},$$

where α, β are the endpoints of γ. It is this quantity that defines a measure on $G(\mathbb{H}^2)$. For details see [Bonahon 1988].

1-4. *Tubular neighborhood about a geodesic.* Suppose ℓ is the vertical half-line rising from the origin in the upper half-space or upper half-plane model. Given $d > 0$, show that the locus of the points of hyperbolic distance d from ℓ consists of the cone of angle ϕ, or the two euclidean lines of angle ϕ from ℓ, where $\sec\phi = \cosh d$.

The corresponding neighborhood about a geodesic which is a semicircle looks like a banana (Figure 1.6).

Next construct a sphere with euclidean center on ℓ which is tangent to the cone of distance d. Find its hyperbolic center which by symmetry also lies on ℓ. *Hint:* Construct the hyperbolic line segment between two opposite points of tangency of the sphere with the cone. It is orthogonal to both the sphere and ℓ. Show that it is a hyperbolic diameter. Denote by $(0, a)$, $(0, b)$ the north and south pole of the sphere with hyperbolic center $(0, c)$. Show that $c^2 = ab$.

Show that in the disk model, the hyperbolic center of a circle coincides with the euclidean center if and only they are at the origin. In the upper half-plane model they never coincide. The corresponding statements hold in three dimensions.

1-5. Let ℓ be a line in the upper half-space model, which we assume without loss of generality to have endpoints $\pm 1/\beta \in \mathbb{R}$ on the boundary plane $\mathbb{C}$. Suppose given two other points on ℓ; their projections to $\mathbb{C}$ lie on the line segment joining the endpoints and so are the form λ/β, μ/β, with $-1 < \lambda < \mu < 1$. Let their heights above $\mathbb{C}$ be s and t, respectively. Show that

$$\chi := \left(-\frac{1}{\beta}, \frac{1}{\beta}, \frac{\lambda}{\beta} + is, \frac{\mu}{\beta} + it\right) = \sqrt{\frac{(1+\lambda)(1-\mu)}{(1-\lambda)(1+\mu)}} < 1. \tag{1.5}$$

Specialize to the case that ℓ is the axis of a loxodromic T with $T(\lambda/\beta + is) = \mu\beta + it$.

1-6. Prove directly (without Gauss–Bonnet) that the angle sum Σ of a hyperbolic triangle Δ satisfies $\Sigma = \pi - \text{Area}\,\Delta$. *Hint:* First prove this for a triangle in the upper half-plane model with a vertex at ∞ and the other two above the points $0, 1 \in \mathbb{R}$. Then show that the area of a general triangle is the difference of the areas of two such ideal triangles. To find the area of the ideal triangle, you can use Green's formula from advanced calculus, plus the fact that the hyperbolic length of the horizontal segment $\{y = t,\ 0 \le x \le 1\}$ goes to 0 as $t \to \infty$.

Go on to prove, as in [Epstein and Marden 1987, A.6.1,2], that the area of a hyperbolic triangle Δ with a side of finite hyperbolic length s is strictly less than s. (*Hint:* The area only increases if the other two sides of Δ have infinite length. Then show, still in the upper half-plane model, that given $a > 0$, the hyperbolic area of the rectangular strip $\{z : 0 < \operatorname{Re} z < s, a < \operatorname{Im} y\}$ is s/a. Use this to show that

$$\left(\frac{dA}{ds}\right)_{s=0} = \sin\theta < 1,$$

where θ is the angle between the short side and one of the vertical sides.)

Deduce that the area of a hyperbolic polygon P with one ideal vertex is less than s, where s is the sum of the lengths of the finite sides. Moreover, the sum of the exterior angles of P is less than $s + 2\pi$.

1-7. Let $T = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ be a Möbius transformation of $\mathbb{C} \cup \infty$ such that $T^2(\infty) \neq \infty$ (so that the isometric circles $\mathcal{I}(T)$ and $\mathcal{I}(T^{-1})$ are distinct). Show that T is the composition of reflection in $\mathcal{I}(T)$, followed by reflection in the perpendicular bisector of the line joining the center of $\mathcal{I}(T)$ to the center of $\mathcal{I}(T^{-1})$, followed by a rotation about the center of $\mathcal{I}(T^{-1})$ of angle ϕ, where $e^{i\phi} = \overline{(a+d)}/(a+d)$. What about the case that $\mathcal{I}(T) = \mathcal{I}(T^{-1})$?

If the trace of T is real, the rotation step is not needed.

Show that every Möbius transformation is the composition of two or four reflections in circles on $\mathbb{S}^2$. (A rotation is the composition of two reflections.) If $\operatorname{tr}(T)$ is real, only two reflections are needed.

1-8. (i) Prove that a Möbius transformation that has a real trace leaves invariant some circle in $\partial\mathbb{H}^3$ (which can be taken as the real line).

(ii) If, in addition, the transformation is loxodromic, it maps every hyperbolic plane that contains its axis onto itself.

(iii) [Van Vleck 1919] Prove that the composition $F = E_2 \circ E_1$ of two elliptics has real trace if and only if there is a circle σ in $\mathbb{S}^2$ containing the fixed points of E_1 and E_2. If the fixed points of E_2 separate the fixed points of E_1 on σ then F is elliptic. If the fixed points do not so separate, then $E_1E_2E_1^{-1}E_2^{-1}$ is loxodromic with real trace.

1-9. Prove the formula for the extension to upper half-space:

$$\left|A\vec{x}_1 - A\vec{x}_2\right| = \frac{1}{|c|^2} \frac{\left|\vec{x}_1 - \vec{x}_2\right|}{\left|\vec{x}_1 - A^{-1}(\infty)\right|\left|\vec{x}_2 - A^{-1}(\infty)\right|}, \qquad c \neq 0, \tag{1.6}$$

and a corresponding formula for $c = 0$. Deduce that the hyperbolic metric is invariant under Möbius transformations and that the designation “isometric hemisphere” is justified as $|A'(\vec{x})| = 1$, where here $|A'|$ denotes the Jacobian determinant.

The extension to upper half-space — in fact to all $\mathbb{R}^3 \cup \infty$ — is conformal. Conversely if $F : D \subset \mathbb{R}^n \to \mathbb{R}^n$ is a conformal mapping, then F is the restriction to D of a Möbius transformation, *provided* $n \geq 3$. This striking result is called *Liouville's Theorem*. Liouville proved it under the assumption that the third partial derivatives of F are continuous; it is now known to be true under much weaker hypotheses on F; see [Vuorinen 1988].

1-10. *Symmetry in isometric circles and planes.* Suppose S preserves the upper half-plane UHP, $S(\infty) \neq \infty$, while T preserves the unit disk $\mathbb{D}$. Prove that the isometric circle $\mathcal{I}(S)$ is characterized by the property that $\mathcal{I}(S) \cap$ UHP is the (hyperbolic) perpendicular bisector of $[i, S^{-1}(i)]$, that is, i and $S^{-1}(i)$ are symmetric about $\mathcal{I}(S)$. Correspondingly prove that $\mathcal{I}(T) \cap \mathbb{D}$ is the perpendicular bisector of $[0, T^{-1}(0)]$, that is, 0 and $T^{-1}(0)$ are symmetric about $\mathcal{I}(T)$.

Deduce that if $A(\mathbb{D}) =$UHP, then A maps $\mathcal{I}(T)$ to $\mathcal{I}(S)$.

Show that the corresponding facts are true for the isometric planes of transformations that preserve the upper half-space and ball models of $\mathbb{H}^3$.

Returning to the upper half-space assertion of Lemma 1.5.4, suppose $p \in \partial\mathbb{H}^3$ is not a fixed point of A (normalized). Given $\vec{x} \in \mathbb{H}^3$, let $e(\vec{x}, A)$ denote the plane which is the perpendicular bisector of the line segment $[\vec{x}, A^{-1}(\vec{x})]$. Then p lies on the circle bounding $e(\vec{x}, A)$ if and only if $\vec{x}$ lies on the plane $e(p, A)$, which has the expression (with $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$)

$$\left| z - \left(p + \frac{(p^2c+pd-pa-b)(\overline{a-pc})}{|a-pc|^2-1} \right) \right|^2 + t^2 = \frac{|p^2c+pd-pa-b|^2}{(|a-pc|^2-1)^2},$$

if $|a - pc| \neq 1$. If $p = \infty$, then $e(p, A)$ reduces to the isometric plane for A. Also the plane $e(\vec{x}, A)$ converges to $e(p, A)$ as $\vec{x} \to p$.

Choose A so that $A^{-1}(\infty) = 0$. Then when $\vec{x} = (0, t)$, the vertical coordinate of $A(\vec{x})$ is $1/(t|c|^2)$. This takes the value $1/|c|$ when $t = 1/|c|$. Thus the horosphere σ at $z = 0$ of euclidean diameter $1/|c|$ is tangent to the horizontal plane P of height

$t = 1/|c|$. Moreover A maps $(0, 1/|c|)$ to $(a/c, 1/|c|)$. Therefore A maps σ onto P. The hemisphere centered at $z = 0$ of radius $1/|c|$ is the isometric circle.

1-11. *Stereographic projection.* Confirm that stereographic projection from the north pole of the unit sphere $\mathbb{S}^2$ to the complex plane $\mathbb{C}$ containing the equator of $\mathbb{S}^2$ is given by the following formulas: If $(x_1, x_2, x_3) \in \mathbb{S}^2$ and $z = x + iy$ is the corresponding point in $\mathbb{C}$ then

$$x = \frac{x_1}{1 - x_3}, \qquad y = \frac{x_2}{1 - x_3}, \qquad |z|^2 = \frac{1 + x_3}{1 - x_3}.$$

This can be extended to map the interior $\mathbb{B}$ of $\mathbb{S}^2$ to upper half-space as follows. Reflection in the unit circle $z \mapsto 1/\bar{z}$ extends to the reflection in $\mathbb{S}^2$ given by $\vec{x} \mapsto \vec{x}/|\vec{x}|^2$. Take however the reflection in the sphere S_N of radius $\sqrt{2}$ about the north pole of $\mathbb{S}^2$:

$$I_1(\vec{x}) = 2\frac{\vec{x} - \vec{k}}{|\vec{x} - \vec{k}|^2} + \vec{k}, \qquad \vec{k} = (0, 0, 1).$$

S_N intersects $\mathbb{S}^2$ in its equator and I_1 pointwise fixes that. Also fixed are the vertical planes through the origin.

Because I_1 sends $(0, 0, 1)$ to ∞, $(1, 0, 0)$ to $(1, 0, 0)$, $(0, 1, 0)$ to $(0, 1, 0)$, and $(0, 0, -1)$ to $(0, 0, 0)$, we see that the image of $\partial\mathbb{B}$ is the plane $\{x_3 = 0\}$ and the image of $\mathbb{B}$ is lower half-space.

Follow I_1 by reflection in the horizontal plane $\{x_3 = 0\}$:

$$I_2 : (x_1, x_2, x_3) \mapsto (x_1, x_2, -x_3).$$

The required extension of stereographic projection is $I = I_2 \circ I_1$. the collection of euclidean half-planes and hemispheres in upper half-space correspond to the collection of spherical caps in $\mathbb{B}$ orthogonal to ∂B. We know this once we know that stereographic projection maps the collection of circles/lines in $\mathbb{C}$ to circles on $\mathbb{S}^2$. From this we can also deduce that I preserves the dihedral angles between intersecting hyperbolic planes.

The group of isometries of $\mathbb{B}$ is then the conjugate of the group of the upper half-space by the Möbius transformation I. The formulas are best found by the method given in the next exercise.

1-12. *Formulas for the ball model.* We follow the elegant treatment presented by Ahlfors [1981]. The notation $x^* = \vec{x}/|x|^2$ for reflection of $\vec{x}$ in the unit sphere will be useful. More generally, given $\vec{a} \in \mathbb{B}^3$, the sphere with center a^* orthogonal to $\partial\mathbb{B}^3$ has radius $(|a^*|^2 - 1)^{1/2}$. The formula for reflection in it is

$$\vec{y} = a^* + (|a^*|^2 - 1)(\vec{x} - a^*)^*.$$

The group of Möbius transformations preserving $\mathbb{B}^3$ is generated by an even number of such reflections. A Möbius transformation that sends $\mathbb{B}^3$ onto itself and a given point $\vec{a}$ to the origin 0 is (dropping the vector notation)

$$T_a x = -a + (1 - |a|^2)(x^* - a)^*.$$

The general transformation that sends $\vec{a} \to 0$ is the composition of T_a followed by a euclidean rotation about 0. The jacobian determinant $|T_a'(x)|$ of T_a represents the local stretch of T_a, the same in all directions. It remains unchanged if T_a is followed by a rotation about 0, so it represents the Jacobian determinant of any element that sends $\vec{a} \to 0$ and preserves $\mathbb{B}^3$.

The Jacobian satisfies

$$|T_a'(x)| = \frac{(1 - |a|^2)}{|x|^2 |x^* - a|^2} = \frac{1 - |a|^2}{\left|x - a\,|x|^2\right|^2}.$$

In addition confirm the formulas

$$|T_a x| = |T_a x - T_a a| = |x - a| \sqrt{T_a'(x)} \sqrt{T_a'(a)},$$

and

$$|T_a x| = \frac{|x - a|}{|a|\,|x - a^*|},$$

so that

$$1 - |T_a x|^2 = \frac{(1 - |x|^2)(1 - |a|^2)}{|a|^2 |x - a^*|^2}.$$

Conclude that

$$\frac{|T_a'(x)|}{1 - |T_a x|^2} = \frac{1}{1 - |x|^2}.$$

In other words, the hyperbolic metric is invariant under any transformation that sends some point $a \to 0$; therefore is invariant under all Möbius transformations preserving the ball.

In fact, for any Möbius transformation T in 3-space, whether or not it preserves a ball or half-space,

$$|T\vec{x} - T\vec{y}| = |T'(\vec{x})|^{1/2} |T'(\vec{y})|^{1/2} |\vec{x} - \vec{y}|.$$

This follows from the fact that any nontrivial Möbius transformation T is the composition of similarity mappings $\vec{x} \mapsto m\vec{x} + \vec{b}$ and the reflection $\vec{x} \mapsto \vec{x}/|x|^2$. The jacobian matrix for the similarity is simply mI. For the reflection it is $(I - 2Q(\vec{x}))/|x|^2$, where the matrix $Q(\vec{x}) = (1/|x|^2)\,(x_i x_j)$ satisfies $Q^2 = Q$ and $(I - 2Q)^2 = I$. In other words, at each point $\vec{x}$, the jacobian $T'(\vec{x})$ is a scalar multiple $|T(\vec{x})|$ of an orthogonal matrix.

1-13. Show the existence of regular n-sided hyperbolic polygons as follows. In the disk model of $\mathbb{H}^2$ start at the origin with a tiny regular n-sided euclidean polygon. Radially expand the polygon insuring by rotational symmetry that all sides remain equal in length. Show that the vertex angle decreases monotonically from the euclidean $(n-2)\pi/n$ to zero when the vertices are on the unit circle. Use the same argument for regular polyhedra.

1-14. If g is a Möbius transformation acting in $\mathbb{B}^3$ with center 0 (really $(0, 0, 0)$) and $d(\cdot, \cdot)$ denotes hyperbolic distance, show that

$$e^{-d(0,g(0))} = \frac{1 - |g(0)|}{1 + |g(0)|}.$$

Also show that for any two points $\vec{x}$, $\vec{y}$ in $\mathbb{B}^3$,

$$\frac{1 - |g(\vec{x})|}{1 - |g(\vec{y})|} \leq 2e^{d(\vec{x},\vec{y})}.$$

Let G be a countable group of Möbius transformations. Show that for $\alpha > 0$,

$$\sum_{g \in G} e^{-\alpha d(0,g(0))} < \infty \quad \text{if and only if} \quad \sum_{g \in G} e^{\alpha d(\vec{x},g(\vec{x}))} < \infty,$$

and correspondingly

$$\sum_{g \in G} (1 - |g(0)|)^\alpha < \infty \quad \text{if and only if} \quad \sum_{g \in G} (1 - |g(\vec{x})|)^\alpha < \infty.$$

Referring back to Exercise 1-12 show that

$$\sum_{g \in G} e^{-\alpha d(0,g(0))} < \infty \quad \text{if and only if} \quad \sum_{g \in G} |g'(0)|^\alpha < \infty.$$

Confirm the analogous formulas for groups acting instead in the unit disk.

1-15. For the group of Möbius transformations acting in the upper half-space model, that is for the simple Lie group $\operatorname{Isom} \mathbb{H}^3$, confirm the *Iwasawa decomposition*

$$\operatorname{Isom} \mathbb{H}^3 = KAN,$$

where K is the compact group of rotations about $(z = 0, t = 1)$, A is the abelian subgroup of loxodromic elements with fixed points $0, \infty$, and N is the nilpotent subgroup of euclidean translations.

1-16. Given four distinct points $z_1, z_2, w_1, w_2 \in \mathbb{C} \cup \infty$, show that there exists a Möbius transformation A such that $A(z_1) = -1$, $A(z_2) = 1$, $A(w_1) = -u$, and $A(w_2) = u$ for some $u \in \mathbb{C}$. Clearly $(z_1, z_2, w_1, w_2) = (-1, 1, -u, u)$. A is uniquely determined if it is required that $|u| \geq 1$. *Hint:* Take $z_1 = -1$, $z_2 = 1$, $w_1 = i$, $w_2 = \zeta$. Find an equation for the coefficients of A. For there to be a nonzero solution, the determinant of the coefficients must vanish.

Consider the hyperbolic lines ℓ with endpoints z_1, z_2 and m with endpoints w_1, w_2. Show that there is a uniquely determined common perpendicular to ℓ and m. Show that the hyperbolic distance between the lines is $\log |u|$.

1-17. Prove that there is a unique largest disk in an ideal triangle and that its hyperbolic radius is $\frac{1}{2}\log 3$. Deduce that any hyperbolic disk in $\mathbb{H}^2$ that meets three mutually disjoint open half-planes must have radius exceeding $\frac{1}{2}\log 3$. (*Hint:* Put $z = 0$ and the vertices of the ideal triangle at equally spaced points on the circle.)

In particular a line in $\mathbb{H}^2$ at hyperbolic distance $\frac{1}{2}\log 3$ from a point $z \in \mathbb{H}^2$ covers exactly one third of the horizon. The union of three hyperbolic lines, each at distance at least $\frac{1}{2}\log 3$ from z and at least one at distance strictly greater than that, do not separate z from $\partial\mathbb{H}^2$.

Conclude that in $\mathbb{H}^3$, a hyperbolic ball B that meets three mutually disjoint open hyperbolic half-spaces must likewise have radius exceeding $\frac{1}{2}\log 3$.

(*Hint:* In the ball model assume that the origin is the center of B. The half-spaces determine three mutually disjoint disks $D_i \subset \partial\mathbb{H}^3$. Denote their spherical radii by r_i. A great circle has length 2π so $\sum 2r_i \leq 2\pi$. For at least one index, $2r_j \leq 2\pi/3$. The distance to the origin of the plane rising from ∂D_j is therefore at least $\frac{1}{2}\log 3$.)

1-18. Given a point $z \in \mathbb{H}^2$ and a geodesic γ not through z show that

$$\sinh d(z, \gamma) = \cot(\theta_z/2),$$

where $d(z, \gamma)$ is the distance from z to γ and $\theta_z < \pi$ is the visual angle γ subtends at z.

Hint: In the disk model take $z = 0$. Then γ is an arc of a circle of euclidean radius say of radius r. Set $a = d(0, \gamma)$. Show that $\sinh a = 1/r$ and $\sinh d(0, a) = 2a/(1-a^2)$. Use the right triangle with euclidean sides of length $1, r, d+r$.

1-19. In the upper half-space model let γ be the geodesic with endpoints ± 1. Show that in the hyperbolic arc length parameter s with basepoint at $z = i$,

$$\gamma(s) = \frac{1 + ie^{-s}}{1 - ie^{-s}} = \tanh s + i \operatorname{sech} s.$$

1-20. Let $\tau = \operatorname{tr} A$ be normalized so that $0 \leq \arg \tau < \pi$. Prove (by putting A in standard form and using induction) that there is a sequence $\{\beta_n\}$, $-\infty < n < \infty$, such that $\beta_0 = 0$, $\beta_1 = 1$,

$$A^n = -\beta_{n-1} I + \beta_n A, \quad \text{and} \quad \beta^{n+1} = -\beta_{n-1} + \tau\,\beta_n.$$

Show that $\beta_{-n} = -\beta_n$. Furthermore, $\beta_n = 0$ for some n if and only if $A^n(z) = \mathrm{id}$.

Set $\tau_n = \operatorname{tr} A^n$, so that $\tau_0 = 2$ and $\tau_1 = \tau$. Then $\tau_{-n} = \tau_n$. Prove that

(i) $\tau_n = -\beta_{n-1} + \beta_{n+1}$;
(ii) if $\tau = \lambda + \lambda^{-1}$, then $\tau_n = \lambda^n + \lambda^{-n}$ and $\beta_n = (\lambda^n - \lambda^{-n})/(\lambda - \lambda^{-1})$;
(iii) $\tau_m\,\tau_n = \tau_{m+n} + \tau_{m-n}$, $\beta_m\,\tau_n = \beta_{m+n} + \beta_{m-n}$, and $\beta_m\,\beta_n = (\tau_{m+n} - \tau_{m-n})/(\tau^2 - 4)$.

Show also that $\lim_{|n|\to\infty} \beta_n = \lim_{|n|\to\infty} \tau_n = \infty$, if $|\tau| > 2$.

Finally show that β_n is a polynomial of degree $|n| - 1$ in τ, and τ_n is of degree $|n|$ in τ. Furthermore

$$\frac{d}{d\tau}\tau_n = n\beta_n, \quad \frac{d}{d\tau}\beta_n = \frac{n\tau_n - \tau\beta_n}{\tau^2 - 4}.$$

The isometric circles $\mathcal{I}(A^{\pm 1})$ are symmetric about the midpoint of line segment joining their centers. Replace A by a conjugate so that the midpoint $A(\infty) + A^{-1}(\infty)$

becomes $z = 0$. Show that A then has the form

$$A = \begin{pmatrix} \frac{1}{2}\tau & \frac{1}{4}(\tau^2 - 4) \\ 1 & \frac{1}{2}\tau \end{pmatrix}. \tag{1.7}$$

Jørgensen [1973] used this form to study the behavior of the cyclic group $\langle A \rangle$ as a function of its trace.

Show that

$$A^n = \begin{pmatrix} \frac{1}{2}\tau_n & \frac{1}{4}(\tau_n^2 - 4)\beta_n^{-1} \\ \beta_n & \frac{1}{2}\tau_n \end{pmatrix}. \tag{1.8}$$

Also $A^k(-z) = -A^{-k}(z)$, for $-\infty < k < \infty$.

1-21. Consider with Tukia [1985c] the 3-manifold

$$K = \{(x_1, x_2, x_3) : x_i \in \mathbb{R} \text{ are distinct and induce the positive orientation.}$$

Define a map $\rho : K \to \text{UHP}$ as follows. Let ℓ be the geodesic in UHP between x_1 and x_2. Let x_3^* be the foot of the perpendicular from x_3 to ℓ. Set $\rho(x_1, x_2, x_3) = x^*$. Prove:

(i) For $z \in \text{UHP}$, the set $\rho^{-1}(z)$ is homeomorphic to the circle $\mathbb{S}^1$, and hence that K is homeomorphic to UHP $\times \mathbb{S}^1$.
(ii) If A is loxodromic and preserves UHP with axis $\ell \in \text{UHP}$, the set $\mathcal{S}(A) := \rho^{-1}(\ell)$ is homeomorphic to $\mathbb{R} \times \mathbb{S}^1$.
(iii) If B is another Möbius transformation preserving UHP, $B(\mathcal{S}(A)) = \mathcal{S}(BAB^{-1})$.
(iv) $\mathcal{S}(A)$ and $B(\mathcal{S}(A))$ are either disjoint, identical, or have intersection $\rho^{-1}(z)$ for some $z \in \text{UHP}$.

Suppose $R = \text{UHP}/G$ is a closed hyperbolic surface. Show that there is a natural discrete action of G on K. Show that K/G is homeomorphic to the unit tangent bundle $T(R)$ of R.

Next, show that any orientation preserving homeomorphism (automorphism) $\alpha : R \to R$ induces an automorphism $\hat{\alpha} : T(R) \to T(R)$ of the 3-manifold $T(R) \equiv K/G$. Moreover, homotopic automorphisms α, α_1 of R correspond to homotopic automorphisms $\hat{\alpha}, \hat{\alpha}_1$ of $T(R)$. This result is attributed to Cheeger and Gromov—see [Casson and Bleiler 1988, pp.54–55] for details. (Hint: Set $x = \rho(x_1, x_2, x_3)$ and $p(x) = (x, \sigma_x)$ where $\sigma_x \subset \ell$ is the oriented segment of length two, centered at x.)

1-22. *Ideal tetrahedra.* On $\partial\mathbb{H}^3$, choose any four distinct points z_1, z_2, z_3, z_4. Then draw the six hyperbolic lines obtained by connecting pairs of points. Each triple of points lies on the edge of a uniquely determined hyperbolic plane. The four hyperbolic planes so obtained pairwise intersect in the six lines. The common exterior of these four planes is a four sided solid called an ideal tetrahedron. It is uniquely determined up to isometry by its four "ideal" vertices z_1, z_2, z_3, z_4.

Now using the upper half-space model, send any one of the vertices to ∞. the three faces meeting at ∞ now become vertical planes. The cross section obtained by intersecting with any sufficiently high horizontal plane $\{t = N\}$ is a euclidean triangle.

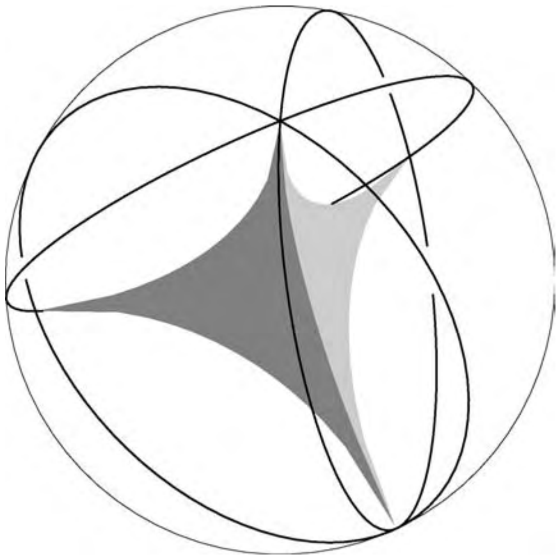


Fig. 1.9. An ideal tetrahedron in the ball model. The dihedral angles are the same as the angles of intersection of the circles on $\partial\mathbb{H}^3$ determined by the faces.

The three angles α, β, γ of that triangle are exactly the dihedral angles formed by the intersection of the corresponding two planes. And of course $\alpha+\beta+\gamma=\pi$.

Label the three other dihedral angles so that δ is opposite β, ϵ is opposite γ, and ρ is opposite α. Any of the four ideal vertices can be sent to ∞. As a consequence the six dihedral angles satisfy four equations. From this, deduce that $\alpha=\rho$, $\beta=\gamma$, $\gamma=\epsilon$. That is, the dihedral angles at opposite edges are the same. In addition, the sum of all the dihedral angles is 2π.

Show that the ideal tetrahedron is uniquely determined by the three angles α, β, γ at the vertex ∞ up to similarity ($z \mapsto az+b$).

Taking thus one of its vertices at ∞ denote the other three ideal vertices by t, u, v, all of which lie in $\mathbb{C}$. These are the vertices of the ideal triangle forming the base of the tetrahedron. Orthogonal projection to $\mathbb{C}$ takes this to a euclidean triangle with vertices t, u, v. Assume the labeling is chosen so that t, u, v in order give the clockwise orientation. Define

$$z_1 = (t, u, v, \infty) = z,$$
$$z_2 = (u, v, t, \infty) = (z-1)/z,$$
$$z_3 = (v, t, u, \infty) = 1/(1-z).$$

Then $z_1 z_2 z_3 = -1$ (which implies $\sum \arg(z_i) = \pi$) and $z_1 z_2 - z_1 + 1 = 0$. Each z_i determines the other two.

Assign the numbers z_1, z_2, z_3 to the vertical edges of the tetrahedron through v, t, u, that is, the three edges at ∞. Now apply a Möbius transformation to the tetrahedron sending a different ideal vertex to ∞ and correspondingly obtain three numbers, using the same clockwise ordering. Show that the same three numbers appear as before and an edge which runs between the original vertex and the new one placed at ∞ is assigned the same number.

Conclude that ideal tetrahedra are uniquely determined up to isometry by three complex numbers z_1, z_2, z_3 that satisfy the two equations above: Starting at any vertex the three edges there are labeled in clockwise order z_1, z_2, z_3. Then the three opposite edges are given the same labeling.

Show that an ideal tetrahedron is also determined up to isometry by the three dihedral angles (which sum to π) along the three edges ending at an ideal vertex. Conversely, given three positive angles which sum to π, there is an ideal tetrahedron with these as dihedral angles at an ideal vertex. The dihedral angles at opposite edges of an ideal tetrahedron are the same.

It may happen that the four ideal vertices lie on a circle in $\mathbb{S}^2$. In this case the "ideal tetrahedron" is degenerate: it lies in a plane. There are three patterns (up to reordering) in which a proper ideal tetrahedron may degenerate, namely the possible orders of the ideal vertices on the circle are (1234), (1342), (1423).

1-23. *Volume of tetrahedra.* Show that the volume of tetrahedra, like the area of triangles, is uniformly bounded above.

An exact formula for the volume of an ideal tetrahedron is derived in [Milnor 1994, §3] and [Ratcliffe 1994, §10.4]. The basic function involved is what Milnor calls the *Lobachevsky function*,

$$Л(\theta) = -\int_0^\theta \log|2\sin u|\,du = \theta\left(1 - \log 2\theta + \sum_1^\infty \frac{2^{2n} B_n}{2n(2n+1)!}\theta^{2n}\right), \tag{1.9}$$

where B_n denotes the n-th Bernoulli number. The series, which is obtained by twice integrating $d^2 Л(\theta)/d\theta^2 = -\cot\theta$, converges for $|\theta| < \pi$ although $Л(\theta)$ itself is periodic with period π. For computations, one generally works with the infinite series. The volume of the ideal tetrahedron with dihedral angles $\alpha, \alpha, \beta, \beta, \gamma, \gamma$ (the opposite dihedral angles of an ideal tetrahedron are equal) is

$$Л(\alpha) + Л(\beta) + Л(\gamma). \tag{1.10}$$

One can also compute the volumes of the regular hyperbolic polyhedra.

Of all hyperbolic tetrahedra, ideal or not, there is a one with the largest volume, which is uniquely determined up to Möbius equivalence [Milnor 1994, p. 200]. It is the ideal tetrahedron whose vertices are the vertices of a regular euclidean tetrahedron inscribed in $\mathbb{S}^2$. All its dihedral angles are $\pi/3$ and its group of orientation preserving hyperbolic symmetries is the group of rotations preserving the euclidean tetrahedron. Its volume is $1.0149\ldots$. (The area of the ideal triangle is π.)

There is a classical variational formula useful in studying deformations of hyperbolic polyhedra. It is called the Schläfli formula [Milnor 1994, p. 281]:

$$dV(P) = -\frac{1}{2}\sum_e L(e)\,d\theta_e.$$

Here P is a hyperbolic polyhedron of volume $V(P)$, the sum is over all edges e of P; $L(e)$ is the length of the edge e and θ_e is the interior dihedral angle along e.

1-24. Recall from Theorem 1.3.1 that if p is a point on a side of a triangle, there is a point on at least one of the other two sides which is of distance at most $\log(1+\sqrt{2}) = \operatorname{arcsinh} 1$ away.

Show that something similar holds for hyperbolic tetrahedra in $\mathbb{H}^3$: there exists $C > 0$ such that if $\vec{x}$ is a point on an edge of the tetrahedron, then the minimum distance from $\vec{x}$ to the union of the other edges does not exceed C. Can you find the optimal C? An analogous property is that there is a constant $C > 0$ such that if $\vec{x}$ is any point in a tetrahedron, the minimum distance of $\vec{x}$ to an edge does not exceed C. See Exercise 1-17.

1-25. We will build two more models of the hyperbolic plane, closely related to one another, and explain their relationship to the conformal models introduced in this chapter. For another approach, see [Cannon et al. 1997, Section 7].

We start with the plane $\mathbb{C} = \mathbb{R}^2$, containing the unit circle S^1 and the real line $\mathbb{R}$, which we complete to $\mathbb{R} \cup \infty$ (the boundary of the upper half-plane model). Let U be the unique Möbius transformation of $\mathbb{C} \cup \infty$ fixing -1 and 1 and taking ∞ to i; you found its expression, $U(z) = (z-i)/(1-iz)$, in Exercise 1-1(iv). When restricted to the real line, U can be thought of as stereographic projection from the point i, that is, it maps $x \in \mathbb{R} \cup \infty$ to

$$U(x) = (\hat{x}_1, \hat{x}_2) = \left(\frac{2x}{x^2+1}, \frac{x^2-1}{x^2+1}\right) \in \mathbb{R}^2 = \mathbb{C};$$

the inverse stereographic projection from the circle to the line (compare Exercise 1-11) is $(\hat{x}_1, \hat{x}_2) \mapsto \hat{x}_1/(1-\hat{x}_2)$.

U conjugates the upper half-plane model and the disk model; in particular, the orientation preserving isometries of the disk model can be thought of as elements of $U\,\mathrm{PSL}(2,\mathbb{R})\,U^{-1} \subset \mathrm{PSL}(2,\mathbb{C})$ — since we know from Exercise 1-1 that the orientation preserving isometries of the upper half-plane model are the elements of $\mathrm{PSL}(2,\mathbb{R})$.

The setup is completed by considering $\mathbb{E}^{1,2}$, which is $\mathbb{R}^3$ with the inner product

$$\langle \vec{x}, \vec{y} \rangle = x_1 y_1 + x_2 y_2 - x_3 y_3. \tag{1.11}$$

(This is studied in detail in exercises 1-26 and 1-27.) The set of vectors in $\mathbb{E}^{1,2}$ having length 0 — that is, satisfying $\langle \vec{x}, \vec{x} \rangle = 0$ — is the *light cone*. The light cone corresponds to the unit circle in $\mathbb{R}^2$ via the usual projectivization map

$$\hat{x}_1 = \frac{x_1}{x_3}, \quad \hat{x}_2 = \frac{x_2}{x_3}.$$

The name "light cone" comes from relativity. Vectors of "imaginary length" ($\langle \vec{x}, \vec{x} \rangle < 0$) are called *timelike*, those lying on the light cone are *lightlike*, and those of positive length are *spacelike*.

Now take a Möbius transformation A preserving $\mathbb{R} \cup \infty$ and having normalized matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$; recall that a, b, c, d are real. We associate to A the linear map $\hat{A}$ of $\mathbb{E}^{1,2}$ given by the matrix

$$\frac{1}{2}\begin{pmatrix} 2(ad+bc) & 2(ac-bd) & 2(ac+bd) \\ 2(ab-cd) & a^2-b^2-c^2+d^2 & a^2+b^2-c^2-d^2 \\ 2(ab+cd) & a^2-b^2+c^2-d^2 & a^2+b^2+c^2+d^2 \end{pmatrix}.$$

Check that:

(i) $\hat{A}$ preserves the inner product of $\mathbb{E}^{1,2}$, and thus leaves the light cone invariant. (*Hint:* Taking the inner product (1.11) of the columns of the matrix of $\hat{A}$ yields zero if the columns are distinct and ± 1 if the columns are the same — the minus sign appearing for the third column only. Thus the matrix is "orthonormal" for the given inner product.)

(ii) The map from the light cone onto itself defined by $\hat{A}$ induces on S^1 a map that coincides with UAU^{-1}.

(iii) $\hat{A}$ also induces a map on the unit disk bounded by S^1 (since the set of timelike vectors in $\mathbb{E}^{1,2}$ is also preserved by $\hat{A}$ and corresponds to the unit disk). This map on the disk takes straight line segments to straight line segments.

The *hyperboloid model* of the hyperbolic plane is the (half-)hyperboloid with equation $x_1^2 + x_2^2 - x_3^2 = -1$, $x_3 > 0$, with the metric induced from the ambient space $\mathbb{E}^{1,2}$. The *projective model* or *Klein model* of the hyperbolic plane is the unit disk in $\mathbb{R}^2$, with the metric transported from the hyperboloid model by the central projection map $(x_1, x_2, x_3) \mapsto (\hat{x}_1, \hat{x}_2)(x_1/x_2, x_1/x_3)$. We now justify these metrics and study the basic properties of these models.

(iv) We first define standard maps from the upper half-plane (UHP) to the hyperboloid and projective disk. Given a point z in the upper half-plane, take any nontrivial $A \in \mathrm{PSL}(2, \mathbb{R})$ that fixes z; show that the corresponding linear map $\hat{A}$ has a timelike 1-dimensional eigenspace in $\mathbb{E}^{1,2}$ that does not depend on the choice of A. We take the intersection of this eigenspace with the hyperboloid as the image of z in the hyperboloid model; likewise we take the projection of this eigenspace onto the projective disk as the image of z in the projective model.

(v) Show that this standard map from UHP to the hyperboloid is an isometry between the hyperbolic metric on UHP and the metric induced on the hyperboloid from the ambient space $\mathbb{E}^{1,2}$. Thus the hyperboloid really is a model of the hyperbolic plane. So is, trivially, the projective disk (since we defined the metric by pullback from the hyperboloid.)

(vi) Show that hyperbolic lines are straight line segments in the projective model. What are they in the hyperboloid model? What are the horocycles in the projective and hyperbolic models?

(vii) The orientation preserving isometries of the hyperboloid and projective models are induced by the linear maps $\hat{A}$, as A ranges over $\mathrm{PSL}(2, \mathbb{R})$. Work out the special cases for A in standard form: $x \mapsto x + 1$, $x \mapsto \lambda^2 x$, and $x \mapsto (x \cos \varphi +$

$\sin\varphi)/(-x\sin\varphi+\cos\varphi)$. For each case show that there is a 1-dimensional fixed eigenspace and determine where it is. Identify what corresponds to the axes and fixed points of loxodromic transformations, and to horocycles for parabolics.

(viii) The map U and the map constructed in (iv) take UHS to the disk model and the projective model, respectively. Composing one with the inverse of the other we get a map that fixes the boundary — see (ii) above. Show that its action of this map on the interior of the disk — that is, the map that conjugates the disk and projective models — is radial and corresponds to stereographic projection onto a hemisphere, followed by orthogonal projection back to the unit circle.

(ix) Consider a hyperbolic polygon in the disk model of $\mathbb{H}^2$ and then, as Poincaré before you, take its counterpart in the projective model. Show that the property that a vertex angle be $<\pi$ is preserved, although the angle itself is not. Recover Poincaré's proof that all of the interior vertex angles of a hyperbolic polygon are $<\pi$ if and only if the polygon is hyperbolically convex.

1-26. [Cannon et al. 1997, p. 66] The space $\mathbb{E}^{1,2}$ of the previous exercise is called *Minkowski space*. To thoroughly understand its metric and that of the hyperboloid model, consider the situation in one dimension lower, taking the hyperbola $x^2-y^2=-1$ in $\mathbb{R}^2=\mathbb{E}^{1,1}$. Suppose $\vec{p(t)}=(x(t),y(t))$ describes the motion of a car on say the upper sheet of the hyperbola. We can express the velocity vector as $\vec{p}'=k(t)(y(t),x(t))$ for a scalar function $k(t)$.

With respect to the inner product $\langle\vec{p}_1,\vec{p}_2\rangle=x_1x_2-y_1y_2$, the vectors $\vec{p}(t),\vec{p}'(t)$ are orthogonal, namely $\langle\vec{p}(t),\vec{p}'(t)\rangle=0$, just as they are with the euclidean metric. But suppose the hyperbolic speed is one: $\langle\vec{p}',\vec{p}'\rangle=1$, in other words that t is the hyperbolic arc length parameter. This forces $|k(t)|=1$. Taking $k=1$ we have the coupled pair of differential equations $x'(t)=y(t)$, $y'(t)=x(t)$. We can solve these by infinite series; in fact for suitable initial values, $y=\cosh t$, $x=\sinh t$.

Deduce that the restriction of the indefinite inner product to the hyperbola gives a definite inner product on tangent vectors or points, specifically in the arclength parameter and distance along the hyperbola,

$$\langle\vec{p}'(t_1),\vec{p}'(t_2)\rangle=\cosh d(\vec{p}(t_1),\vec{p}(t_2))=-\langle\vec{p}(t_1),\vec{p}(t_2)\rangle.$$

1-27. We now consider Minkowski space in arbitrary dimension. The inner product in $\mathbb{E}^{1,n}$ is

$$\langle\vec{x},\vec{y}\rangle=\sum_1^n x_iy_i-x_{n+1}y_{n+1},$$

where $\vec{x}=(x_1,\dots,x_{n+1})$. Two nonzero vectors are called orthogonal if $\langle\vec{x},\vec{y}\rangle=0$.

A vector $\vec{x}$ is called *timelike, lightlike, or spacelike* according to whether $|\vec{x}|^2=\langle\vec{x},\vec{x}\rangle$ is negative, zero, or positive. The collection of lightlike vectors forms the *light cone* $\{\vec{x}:|\vec{x}|=0\}$; the upper sheet of the cone is denoted by L^+; it is asymptotic to the upper-sheet hyperboloid $\mathcal{H}^n=\{\vec{x}\in\mathbb{R}^{n+1}:|\vec{x}|^2=-1,\ x_{n+1}>0\}$.

A ray from the origin in L^+ corresponds to a point on $\partial\mathbb{H}^n=\mathbb{S}^{n-1}$.

An n-dimensional hyperplane P with a *spacelike* normal vector $\vec{v}$, namely $P = \{\vec{x} \in \mathbb{E}^{1,n} : \langle \vec{x} - \vec{\zeta}, \vec{v} \rangle = 0\}$ for some $\vec{\zeta} \in P$, intersects $\mathcal{H}^n$ in a $(n-1)$-dimensional hyperbolic subspace.

A plane P with a *timelike* normal vector $\vec{v} \neq 0$ intersects $\mathcal{H}^n$ in a (possibly degenerate) $(n-1)$-sphere.

For the borderline case, a plane P with a *lightlike* normal vector $\vec{v} \neq 0 \in L^+$ intersects $\mathcal{H}^n$ in an $(n-1)$-dimensional horosphere. There is a unique lightlike vector v such that $P = \{\vec{x} \in E^{1,n} : \langle \vec{x}, \vec{v} \rangle = -1\}$. The ray of L^+ from the origin through $\vec{v}$ corresponds to the point at ∞ of the horosphere. The corresponding horoball is

$$\{\vec{x} \in \mathbb{H}^n : 0 \geq \langle \vec{x}, \vec{v} \rangle \geq -1\}.$$

As $\vec{v}$ increases along its ray, the horoball contracts to its point at ∞. Because of the peculiarities of the metric, the normal vector is simultaneously orthogonal to and parallel to the plane P.

Now set $n = 3$. $\mathbb{E}^{1,3}$ is the space/time of relativity theory. Parabolic, elliptic, and loxodromic Möbius transformations correspond to linear transformations conjugate to the respective linear isometries:

$$\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & \frac{3}{2} & -\frac{1}{2} \\ 1 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cosh\lambda & \sinh\lambda \\ 0 & 0 & \sinh\lambda & \cosh\lambda \end{pmatrix}.$$

A parabolic transformation has only one eigenvalue, which is 1, and preserves a single ray in L^+. A loxodromic has two eigenvalues not on the unit circle, which are $\cosh\lambda \pm \sinh\lambda$ (its other two are $e^{\pm i\theta}$), and preserves two rays in L^+. All the eigenvalues of an elliptic lie on the unit circle; the eigenvalue 1 is repeated twice, the other two are $e^{\pm i\theta}$.

Show that if $x \mapsto Ax$ is parabolic with fixed point $u \in \mathbb{R} \cup \infty$, the corresponding linear transformation fixes every point on the corresponding ray of the light cone L^+ in $\mathbb{E}^{1,2}$:

$$x_1 = \frac{2u}{u^2+1} x_3, \qquad x_2 = \frac{u^2-1}{u^2+1} x_3.$$

If instead $x \mapsto Ax$ is loxodromic with fixed points p, q, show that the corresponding linear transformation fixes each point of the ray from $(0, 0, 0,) \in L^+$ orthogonal to the plane spanned by the two rays in the light cone determined by the fixed points. There is expansion by the factor ρ^2 along the ray for the attracting fixed point, and contraction by ρ^{-2} along the ray for the repelling one. In fact, it is best to work out first the geometry for $\mathbb{E}^{1,2}$. For more detail see [Hodgson and Weeks 1994].

Explain why elliptic, parabolic, and hyperbolic transformations of $\mathbb{H}^2$ are associated with ellipses, parabolas, and hyperbolas, respectively. (Each has an invariant plane that cuts the light cone in the respective conics.)

Here is an alternate way to view $\mathcal{H}^2$. Associate column vectors in $\mathbb{E}^{1,2}$ with real symmetric matrices as follows:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longleftrightarrow \begin{pmatrix} x_3 - x_2 & x_1 \\ x_1 & x_3 + x_2 \end{pmatrix}.$$

Again let A be a normalized 2×2 matrix with real coefficients. Consider the action

$$\begin{pmatrix} x_3 - x_2 & x_1 \\ x_1 & x_3 + x_2 \end{pmatrix} \mapsto A \begin{pmatrix} x_3 - x_2 & x_1 \\ x_1 & x_3 + x_2 \end{pmatrix} A^t,$$

which preserves the determinant $-(x_1^2 + x_2^2 - x_3^2)$. Relate this action to the action of $\hat{A}$ in Exercise 1-25.

Passing on to $\mathbb{E}^{1,3}$, associate vectors with hermitian matrices as follows:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \longleftrightarrow \begin{pmatrix} x_4 - x_1 & x_2 + i x_3 \\ x_2 - i x_3 & x_4 + x_1 \end{pmatrix}.$$

Let A be a normalized 2×2 matrix now with complex coefficients. It acts on $\mathbb{E}^{1,3}$ by

$$\begin{pmatrix} x_4 - x_1 & x_2 + i x_3 \\ x_2 - i x_3 & x_4 + x_1 \end{pmatrix} \mapsto A \begin{pmatrix} x_4 - x_1 & x_2 + i x_3 \\ x_2 - i x_3 & x_4 + x_1 \end{pmatrix} \bar{A}^t,$$

which leaves the determinant $-(x_1^2 + x_2^2 + x_3^2 - x_4^2)$ invariant. Show that the action preserves $\mathcal{H}$ (the upper sheet of the hyperboloid). Show that this is the action brought over from $\mathbb{C} \cup \infty$ to $\mathbb{E}^{1,3}$ by stereographic projection and homogeneous coordinates.

See [Weeks 1993] or [Greenberg 1962] for more details.

1-28. *The quaternion description.* Upper half-space can be neatly described by the division ring of quaternions. Quaternions can be identified with the group of matrices

$$\mathcal{Q} = \left\{ \begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix} : u, v \in \mathbb{C} \right\}$$

as follows. Set

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{i} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \boldsymbol{j} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \boldsymbol{k} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Then $\boldsymbol{i}\boldsymbol{j} = \boldsymbol{k}$, $\boldsymbol{j}\boldsymbol{k} = \boldsymbol{i}$, $\boldsymbol{k}\boldsymbol{i} = \boldsymbol{j}$. Also $\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -\mathbf{1}$. Writing $u = u_1 + i u_2$, $v = v_1 + i v_2$, set

$$\begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix} = u_1 + u_2 \boldsymbol{i} + v_1 \boldsymbol{j} + v_2 \boldsymbol{k} = u + v \boldsymbol{j}.$$

For $\boldsymbol{z} = u + v\boldsymbol{j} \in \mathcal{Q}$, define $\bar{\boldsymbol{z}} = \bar{u} - v\boldsymbol{j}$, and $|\boldsymbol{z}|$ by $|\boldsymbol{z}|^2 = |u|^2 + |v|^2$. Note that for $c \in \mathbb{C}$, $c\boldsymbol{j} = \boldsymbol{j}\bar{c}$.

Points in $\mathbb{R}^3$ can now be denoted as the special quaternion $\boldsymbol{z} = z + t\boldsymbol{j}$, $z \in \mathbb{C}$, $t \in \mathbb{R}$.

Given $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, with $ad - bc = 1$, $a, b, c, d \in \mathbb{C}$, show that the action of the corresponding Möbius transformation in upper half-space is described by

$$A(z) = (az + b)(cz + d)^{-1} = (zc + d)^{-1}(za + b),$$

where the quaternion $A(z)$ is of the same special type as z. For details see [Ahlfors 1981], [Beardon 1983] and [Fenchel 1989].

1-29. Let $T : \mathcal{E}^{1,n} \to \mathcal{E}^{1,n}$ be a (not necessarily orientation preserving) linear transformation that maps the (upper-sheet) hyperboloid $\mathcal{H}^n$ onto itself and preserves the inner product:

$$\langle Tx, Ty \rangle = \langle x, y \rangle.$$

Prove the following necessary and sufficient condition for T to be of this type, where M is the matrix corresponding to T:

$$M^t J M = J \quad \text{and} \quad M_{n+1,n+1} > 0.$$

Then prove that T acts isometrically on $\mathcal{H}^n$, and that the only isometries of $\mathcal{H}^n$ are of this type. See [Cannon et al. 1997] if you need help.

1-30. *Relation of hyperbolic and euclidean metrics.* Suppose $f : \mathbb{D} \to \Omega$ is a conformal mapping from the unit disk. The *hyperbolic* or *Poincaré* metric in Ω is defined to be

$$\rho(w)\,|dw| = \frac{2|f'(z)|\,|dz|}{1 - |f(z)|^2}.$$

The Riemann map f is uniquely determined only up to postcomposition by Möbius transformations, but such compositions do not change the metric in Ω. Prove using the Koebe $\frac{1}{4}$-Theorem from conformal mapping theory (see [Pommerenke 1992, 4.6(6)]) that

$$\frac{1}{2d(w, \partial\Omega)} \le \rho(w) \le \frac{2}{d(w, \partial\Omega)},$$

where $d(w, \partial\Omega)$ is the shortest euclidean distance from w to the boundary of Ω. Equality on the left holds if and only if the complement of Ω is a ray to ∞ in a line through w. Equality on the right holds if and only if Ω is a round disk centered at w.

Suppose instead that $\Omega \subset \mathbb{C}$ is a multiply, possibly infinitely, connected region whose complement contains at least two points (plus ∞). That is, assume that Ω carries a hyperbolic metric $\rho(w)\,|dw|$ that arises from projection from the unit disk, its universal cover (see Chapter 2). According to [Beardon and Pommerenke 1978], [Pommerenke 1984] there exists $C = C(\Omega) > 0$ such that

$$\frac{C}{d(w, \partial\Omega)} \le \rho(w) \le \frac{2}{d(w, \partial\Omega)}, \tag{1.12}$$

if and only if $X = \partial\Omega$ has the property called uniformly perfect.

The notion of uniformly perfect is really directed to multiply connected regions as when Ω is simply connected $\neq \mathbb{C}$ we already know that Equation 1.12 holds with $C =$

$1/2$. Yet it simplifies terminology to simply declare that closed, connected sets $X \in \mathbb{S}^2$ with more than two points are automatically uniformly perfect (their complementary components are simply connected).

So consider closed sets $X \subset \mathbb{S}^2$ that are *not* connected. Then $X \subset \mathbb{S}^2$ is called *uniformly perfect* [Beardon and Pommerenke 1978; Pommerenke 1984] if there exists a constant $M < \infty$ such that any annular region $A \subset \mathbb{S}^2 \setminus X$ that separates the components of X has modulus $\mathrm{mod}(A) < M$. Here $\mathrm{mod}(A) = (\log r)/2\pi$ where $f : A \to \{1 < |z| < r\}$ is a conformal map. The uniformly perfect condition relates to a requirement that X be "uniformly thick" at each of its points, independent of scaling by Möbius transformations. For example if X contains an isolated point, there would be a separating annular region of arbitrarily large modulus.

If X is uniformly perfect, the boundary of any multiply connected complementary component is uniformly perfect as well.

An equivalent condition is that X is uniformly perfect if for every $a, b \in X$ and $w \notin X$ there exists $c \in X$ such that for some constant M the cross ratio satisfies

$$\frac{1}{M} \le |(a, b, c, w)| \le M.$$

Most importantly for us, the limit set $\Lambda(G)$ of any finitely generated (nonelementary) kleinian group G is uniformly perfect [Pommerenke 1984]! Therefore each component Ω of the complement $\Omega(G)$ satisfies Equation 1.12 for some $C = C(\Omega)$. On the other hand, there exist infinitely generated Schottky groups whose limit sets are not uniformly perfect. See Chapter 2 for the basic properties of kleinian groups.

1-31. *Anti-Möbius transformations.* An anti-Möbius transformation A can be expressed as $A = B \circ J$, where J is complex conjugation and B is orientation preserving. Show that in $\mathbb{S}^2$, A either pointwise fixes a circle, or it has zero, one or two fixed points. Examples:

$$z \mapsto \frac{1}{\bar{z}}, \quad z \mapsto -\frac{1}{\bar{z}} + 1, \quad z \mapsto \bar{z} + 1, \quad z \mapsto 2\bar{z}.$$

The extension of A to open upper half-space pointwise fixes a plane, a line, or a point.

A finer classification [Fenchel 1989, pp. 48–53] has the possibilities grouped in three conjugacy (by an orientation preserving transformation) classes. The first two are called *involutive* since the elements have order two. The third conjugacy class consists of elements T with the property that T^2 is loxodromic or parabolic. These arise as described below.

To understand the classification, we have to anticipate the result of Lemma 7.3.1 that a loxodromic or parabolic transformation has a square root which is a Möbius transformation of the same type. From Lemma 7.1.2 we take the result that any Möbius transformation that interchanges two distinct points in $\mathbb{S}^2$ is elliptic of order two. Fenchel's classification is as follows:

Reflection in a plane: Reflection J in a plane in $\mathbb{H}^3$.

Reflection in a point $c \in \mathbb{H}^3$: Suppose $c = (0, 0, t)$ in the upper half-space model of $\mathbb{H}^3$. Let ℓ denote the vertical axis and P be a plane through c and orthogonal to ℓ. The point-reflection in c has the form $J \circ E$ where E is the elliptic of order two $E : (x, y, t) \mapsto (-x, -y, t)$ with rotation axis ℓ and J is reflection in P.

Noninvolutive anti-Möbius transformations: Let T^2 be loxodromic or parabolic and pick a square root T. Choose any $x \in \mathbb{S}^2$ distinct from its fixed points and set $y = T^{-1}(x), z = T(x)$ so that the three points x, y, z are distinct. Let $\tau \subset \mathbb{H}^3$ be the line with endpoints y, z and let P be the plane orthogonal to τ whose boundary passes through x. Let J denote reflection in P. Then $T \circ J = E$ interchanges x, z so that it is elliptic of two with rotation axis $\ell \subset P$. Thus $T = E \circ J$.

Show that an elliptic of order two is itself the composition of reflections in two orthogonal planes intersecting in its rotation axis.

Given two distinct circles $C_1, C_2 \in \mathbb{S}^2$, show that there is a circle C^* such that reflection in C^* interchanges C_1 and C_2.

1-32. *Hilbert's metric.* Let Ω be a bounded, euclidean convex domain in $\mathbb{R}^n$. Then every euclidean straight line that contains a point of Ω intersects its boundary $\partial\Omega$ in exactly two points. Given two points $x, y \in \Omega$, denote by $x', y' \in \partial\Omega$ the two points of intersection of the line L through x, y with $\partial\Omega$, so labeled that x separates x' from y along L. Consider the expression

$$d(x, y) = \log \frac{|x - y'|\,|y - x'|}{|x - x'|\,|y - y'|}.$$

Show that Ω is a complete metric space with metric $d(\cdot\,,\cdot)$. The geodesics in this space are the euclidean line segments.

An affine map of R^n sends the metric to the Hilbert metric of the image domain. It turns out that $d(\cdot, \cdot)$ is a riemannian metric on Ω if and only if Ω is an ellipsoid and d is the hyperbolic metric on Ω, here considering Ω as the Klein model of $\mathbb{H}^n$.

Show that if $\partial\Omega$ contains a straight line segment, then two rays from any point $O \in \Omega$ to different points on the line are of uniformly bounded distance apart.

1-33. *Hyperbolic curvature of arcs.* Suppose $z = z(t) = x(t) + iy(t)$ is a parametrized arc in $\mathbb{H}^2$, $-m < t < m$, with continuous second derivative. In euclidean geometry, the parameter t is an *arc length* parameter if and only if $|z'(t)| = 1$. In hyperbolic geometry, in the upper half-plane model, t is arc length parameter if and only if

$$\frac{|z'(t)|}{y(t)} = 1, \quad -m < t < m.$$

In euclidean geometry with arc length parameter s the curvature of $z = z(s)$ is defined as

$$\kappa_e(s) = \lim_{\Delta s \to 0} \frac{|\Delta\varphi|}{|\Delta s|}.$$

Here $\Delta\varphi = \varphi(s+\Delta s) - \varphi(s)$, where the slope of the tangent line at the point $z(s)$ is $\tan\varphi(s)$ and correspondingly $\tan\varphi(s+\Delta s)$ at $z(s+\Delta s)$.

In hyperbolic geometry, the curvature is defined by exactly same formula but the meaning of the terms is hyperbolic: s is the hyperbolic arc length parameter and $\Delta\varphi$ is the angle between the *hyperbolic tangent lines* at $z(s)$ and $z(s+\Delta s)$.

Another characterization valid in both the euclidean and hyperbolic situations is this: About the point $s=0$ take the local coordinate system determined by the tangent vector $\vec{\alpha}$ to the curve at $s=0$ and the normal $\vec{\beta}$ to the curve at that point. In this new coordinate system the curve $\zeta = \zeta(s)$ has the expansion

$$\zeta(s) = s\vec{\alpha} + \kappa(0)\frac{s^2}{2}\vec{\beta} + O(s^3).$$

Using Exercise 1-19 confirm that the curvature of a geodesic is zero.

Curvature of a horocycle. In the upper half-plane model, consider the horocycle $z(t) = ai + t$, $a > 0$ (which is already in euclidean arc length parameter). In hyperbolic arc length parameter, the equation is

$$z(s) = ai + as.$$

Along the horocycle, $a\Delta s = \Delta x$. So it suffices to compute the limit of $\frac{a\Delta\varphi(s)}{\Delta x}$. Actually it suffices to take the case $a = 1$.

The hyperbolic tangent lines to $\{z : \operatorname{Im} z = 1\}$ are the semicircles with center on $\mathbb{R}$ and unit radius. Because the horocycle is invariant under the continuous group of translations, it suffices to make the computation at say $s = 0$. To do this take the unit semicircle centered at $z = 0$ and the semicircle centered at $z = \Delta x$. Find their point of intersection, and then find the angle between them at this point (choose the angle so that it would be zero if the two tangents coincided); this is our $\Delta\varphi$. Taking the limit as $\Delta x \to 0$ we find (in sharp contrast to the euclidean case) that

$$\text{curvature of a horocycle} = 1.$$

Curvature of an equidistant arc. In the upper half-plane model measure distances from the vertical half line. Consider the line $y = cx$, $c > 0$, making angle θ with the vertical. In its hyperbolic arc length parameter,

$$z(s) = e^{s\cos\theta} e^{i(\pi/2-\theta)},$$

where $c = \cot\theta$. The line is invariant under the continuous group $z \to kz$, $k > 0$, so it suffices to make the computation at (m, mc) where $m = \sin\theta$. Find the angle of intersection $\Delta\varphi$ between the hyperbolic tangent lines at (m, cm) and $e^{s\cos\theta}(m, cm)$. Show that at the point $s = 0$,

$$\lim_{\Delta s\to 0}\left|\frac{\Delta\varphi}{\Delta s}\right| = \frac{\cos\theta}{c}.$$

Conclude that

$$\text{curvature of a line of distance } d \text{ from a geodesic} = \tanh d = \sin\theta. \tag{1.13}$$

For example, the 45° line has curvature $1/\sqrt{2}$.

Curvature of a circle. In the unit disk model, consider a circle of euclidean radius R about the origin. In hyperbolic arc length coordinates starting from $(R, 0)$ its equation is

$$z(s) = re^{is(1-R^2)/(2R)}.$$

Take the geodesics tangent to the circle at $z = R$ and at $z = Re^{i\theta}$. Find their point of intersection (x_0, y_0) within the unit disk and then their angle $\varphi = \varphi(\theta)$ of intersection. After a long calculation, for example by calculating $\lim_{\theta\to 0} \varphi/\theta$, one finds

$$\left.\frac{d\varphi}{d\theta}\right|_{\theta=0} = \frac{1+R^2}{1-R^2}.$$

Because the circle is invariant under the continuous group of rotations, it suffices to make this calculation at a single point.

Note that

$$\frac{d\varphi}{ds} = \frac{1-R^2}{2R}\frac{d\varphi}{d\theta}.$$

The hyperbolic radius ρ satisfies $e^\rho = (1+R)/(1-R)$ so that $\coth\rho = (1+R^2)/2R$. We end up with the following formula (compare Section 1.4):

$$\text{curvature of a circle of hyperbolic radius } \rho = \coth\rho. \tag{1.14}$$

In both 1.13 and 1.14, the curvature approaches 1 as $R \to 1$ or $\rho \to \infty$. Why?

Summary:

$$\text{constant curvature}\begin{cases} <1 \iff \text{curves of finite distance from a geodesic,}\\ =1 \iff \text{horocycles,}\\ >1 \iff \text{circles.}\end{cases}$$

In $\mathbb{H}^3$ consider a surface of distance d from a hyperbolic plane. The nearest point map that projects the surface to the plane scales hyperbolic distances by a factor $1/\cosh d$. Conclude from this that:

$$\text{Gaussian curvature of a surface of distance } d \text{ from a plane} = -\operatorname{sech}^2 d.$$

Also deduce:

$$\text{Gaussian curvature of surface of distance } d \text{ from a line} = 0.$$

1-34. *Conjugation by involution: Wada's Lemma* [2003]. Suppose A, B are two Möbius transformations which do not share a fixed point. If A and B are conjugate, show that there is a Möbius transformation Q, with $Q^2 = \mathrm{id}$, such that $A = QBQ$. That is, show that A and B are conjugate by an involution.

Hint: For the parabolic case take $A = \left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$. If $B = XAX^{-1}$, set $Y = XT$ where $T = \left(\begin{smallmatrix}1&t\\0&1\end{smallmatrix}\right)$. Note that $YAY^{-1} = B$.

1-35. *Parametrization of two generator groups with parabolic commutator* [Jørgensen 2003]. Assume that the matrices X, Y satisfy the relation $[X, Y^{-1}] = \left(\begin{smallmatrix} -1 & -2 \\ 0 & -1 \end{smallmatrix}\right)$. Introduce the notation

$$a = \operatorname{tr} X, \quad b = \operatorname{tr} Y, \quad c = \operatorname{tr} XY^{-1} = \operatorname{tr} Y^{-1}X, \quad c' = \operatorname{tr} XY.$$

Applying Lemma 1.5.6, these quantities are connected by the relations

$$abc = a^2 + b^2 + c^2, \quad cc' = a^2 + b^2.$$

These relations are symmetric with respect to c, c', that is, $abc' = a^2+b^2+c^2$. If one of the traces a, b, c is zero, that element represents an elliptic of order two. If two of the traces are zero, say $a = b = 0$, then $c = 0$. If none of the traces are zero,

$$\frac{a}{bc} + \frac{b}{ac} + \frac{c}{ab} = 1 = \frac{a}{bc'} + \frac{b}{ac'} + \frac{c'}{ab}. \tag{1.15}$$

Complex numbers u, v, w that lie on the hyperplane

$$\mathfrak{P} = \{(u, v, w) \in \mathbb{C}^3 : u + v + w = 1\}$$

are called *complex probabilities*. From our perspective, the *singular subset* $\Sigma_0 \in \mathfrak{P}$ is the union of the three coordinate lines:

$$\Sigma_0 = \{(u, v, w),\ u + v + w = 1 : u = 0 \text{ or } v = 0 \text{ or } w = 0\}.$$

Note that $\mathfrak{P} \setminus \Sigma_0$ is connected.

We are now ready to parametrize two generator groups with parabolic commutator by complex probabilities. Confirm the following facts.

Suppose $u, v, w \in \mathbb{C}$ are nonvanishing numbers such that $u + v + w = 1$. Set

$$d = \frac{1}{uvw}, \quad a = \sqrt{ud}, \quad b = \sqrt{vd},$$

where the arguments are chosen so that $-\pi < \arg d,\ \arg a,\ \arg b \le \pi$. Choose $\arg c$ so that

$$c = \sqrt{wd}$$

satisfies

$$abc = \sqrt{ud}\sqrt{vd}\sqrt{wd} = d.$$

Set

$$X = \begin{pmatrix} a - b/c & a/c^2 \\ a & b/c \end{pmatrix}, \quad Y^{-1} = \begin{pmatrix} b - a/c & -b/c^2 \\ -b & a/c \end{pmatrix}. \tag{1.16}$$

Then

$$XY^{-1}X^{-1}Y = \begin{pmatrix} -1 & -2 \\ 0 & -1 \end{pmatrix}, \quad XY^{-1} = \begin{pmatrix} c & -1/c \\ c & 0 \end{pmatrix}, \quad Y^{-1}X = \begin{pmatrix} c & 1/c \\ -c & 0 \end{pmatrix},$$

and $\operatorname{tr} X = a$, $\operatorname{tr} Y = b$, and $\operatorname{tr} XY^{-1} = c$ with $a, b, c \neq 0$. Conversely, if X and Y satisfy $[X, Y^{-1}] = \left(\begin{smallmatrix} -1 & -2 \\ 0 & -1 \end{smallmatrix}\right)$, and the numbers $a = \operatorname{tr} X$, $b = \operatorname{tr} Y$ and $c = \operatorname{tr} XY^{-1}$ are nonzero, and $XY^{-1}(0)=\infty$, then X and Y^{-1} have the matrix representations (1.16). In summary,

$$\operatorname{tr} X = a, \quad \operatorname{tr} Y = b, \quad \operatorname{tr} XY^{-1} = c,$$

$$u=\frac{a}{bc}, \quad v=\frac{b}{ac}, \quad w=\frac{c}{ab}, \quad a^2=\frac{1}{vw}, \quad b^2=\frac{1}{uw}, \quad c^2=\frac{1}{uv}, \quad abc=\frac{1}{uvw},$$

$$a^2+b^2+c^2=abc, \quad u+v+w=1.$$

And in the opposite direction:

Lemma 1.6.1. *Set* $K=\left(\begin{smallmatrix}-1 & -2\\ 0 & -1\end{smallmatrix}\right)$. *Suppose the matrices* X, Y *satisfy*

$$XY^{-1}X^{-1}Y \;=\; K,$$

and the transformation corresponding to X *does not fix* ∞. *Then* X *and* Y^{-1} *have the form*

$$X \;=\; \begin{pmatrix} * & * \\ \tau_X & * \end{pmatrix}, \quad Y^{-1} \;=\; \begin{pmatrix} * & * \\ -\tau_Y & * \end{pmatrix}.$$

Proof. As usual, all matrices have determinant one. Set

$$X \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad Y^{-1} \;=\; \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}.$$

Replace X and Y^{-1} by the conjugates WXW^{-1} and $WY^{-1}W^{-1}$, where $W=\left(\begin{smallmatrix}1 & d/c\\ 0 & 1\end{smallmatrix}\right)$, so that $WXW^{-1}(0)=\infty$; conjugation by the translation W leaves K unchanged and it also leaves the entries c and γ in the matrix for X and Y^{-1} unchanged. The identity $XY^{-1}=KY^{-1}X$ gives the four equations

$$c\beta=-b\gamma, \quad c\alpha=-a\gamma-c\delta, \quad a\alpha+b\gamma=-a\alpha-c\beta-2a\gamma-2c\delta,$$
$$a\beta+b\delta=-b\alpha-2b\gamma.$$

Substituting the first and second into the third we wind up with either $c=a=\tau_X$ or $\alpha=0$. In the former case, $-\gamma=\alpha+\delta=\tau_Y$ and $\beta=-\tau_Y/\tau_X^2$, since $bc=-1$. With the normalization $X(0)=\infty$, leaving aside the formulas for α, δ, we have

$$X \;=\; \begin{pmatrix} \tau_X & -1/\tau_X \\ \tau_X & 0 \end{pmatrix} \quad Y^{-1} \;=\; \begin{pmatrix} \alpha & -\tau_Y/\tau_X^2 \\ -\tau_Y & \delta \end{pmatrix}. \tag{1.17}$$

Suppose instead $\alpha=0$. Using the fact $bc=-1$ and $\beta\gamma=-1$, we find from the first equation that $c^2=-\gamma^2$, or $\gamma=\pm ci$. The second equation becomes $a\gamma+c\delta=0$, or $\delta=\mp ai$; in other words $\tau_Y=\mp\tau_X i$. The fourth equation becomes $a\beta+b\delta=-2b\gamma$ and upon rewriting in terms of c and a yields $c=a=\tau_X$. Putting it all together we actually have a special case of (1.17),

$$X \;=\; \begin{pmatrix} \tau_X & -1/\tau_X \\ \tau_X & 0 \end{pmatrix}, \quad Y^{-1} \;=\; \begin{pmatrix} 0 & 1/\tau_Y \\ -\tau_Y & \tau_Y \end{pmatrix}.$$

In this case $Y^{-1}X$ and XY^{-1} represent the order two elliptics $z\mapsto -z$ and $z\mapsto -z+2$. $\square$

The bottom line is that given the traces of generators X, Y, plus the trace of XY^{-1}, plus the condition that the trace of the commutator is -2, the group is uniquely determined up to conjugation. Moreover, by the trace identities, the trace of any element of $\langle X, Y\rangle$ is a polynomial in the initial traces. The systematic method for doing this uses the modular diagram and the Farey sequence; it is spelled out in [Mumford et al. 2002].

1-36. Suppose the Möbius transformations A, B are such that $[A, B]$ is parabolic. Assume that $A, B, C = BA$ all have real traces. Show that for some Möbius T, the entries of the normalized matrices for TAT^{-1}, TBT^{-1} are real. Consequently the group $T\langle A, B\rangle T^{-1}$ preserves the upper half-plane. What if $[A, B]$ is instead elliptic? Show by modifying Jørgensen's method of complex probabilities that the same conclusion holds.

1-37. *Two-jets of locally injective analytic functions.* Suppose $f(z)$ is a locally injective analytic function (that is, $f'(z) \neq 0$) in a neighborhood of $z_0 \in \mathbb{C}$. Show that there is a Möbius transformation $M(f; z_0)$ uniquely determined by the three properties:

$$M(f; z_0)(z_0) = f(z_0), \quad M(f; z_0)'(z_0) = f'(z_0), \quad M(f; z_0)''(z_0) = f''(z_0).$$

The value of the first two derivatives of f at z_0 is called the *two-jet* of f at z_0. If A is any Möbius transformation, $M(Af; z_0) = AM(f; z_0)$.

Thurston [1986d] showed that for any $v \neq 0 \in \mathbb{C}$,

$$vS_f(z_0) = v\left(\left(\frac{f''}{f}\right)' - \frac{1}{2}\left(\frac{f''}{f'}\right)^2\right)_{z_0} = \frac{\partial^2}{\partial z^2} M(f; z)^{-1} M(f; z + tv)\Big|_{t=0, z=z_0}.$$

Here $S_f(z_0)$ is called the *schwarzian derivative* (Exercise 6-8).

Suppose now that $f : \mathrm{UHP} \to \mathbb{C}$ is a conformal map. There is an extension F of f to $\mathbb{H}^3$, taken as UHS, determined as follows.

Denote by $P \subset \mathbb{H}^3$ the vertical half-plane rising from $\mathbb{R}$. Identify P with UHP by orthogonal projection. Given $x \in \mathbb{H}^3$, let ℓ_x be the geodesic through x that is orthogonal to P. Denote by $r(x)$ its intersection with $P \equiv$UHP.

Set

$$F(x) = M(f; r(x)).$$

Show that F is continuous. It is also equivariant if f is so: Suppose G is a fuchsian group acting on UHP, and on $\mathbb{H}^3$, and f satisfies for all $\gamma \in G$, $f \circ \gamma = \phi(\gamma) \circ f$, where ϕ is a homomorphism of G to another group, not necessarily discrete. Then

$$M(f; r(\gamma x)) = \phi(\gamma) M(f; r(x))\gamma^{-1},$$

and consequently $F \circ \gamma = \phi(\gamma) \circ F$. Although F is not necessarily a local homeomorphism, there exists $d > 0$ such that F is a local homeomorphism outside of a distance-d neighborhood of P. See [Bromberg 2000] for details.

1-38. *The hyperbolic Gauss map.* Suppose $S \subset \mathbb{H}^3$ is a smoothly immersed, *oriented* surface in the ball model. Given $\zeta \in S$ let $\vec{n}^{\zeta}$ denote the geodesic ray normal to S at ζ

and denote its endpoint on $\mathbb{S}^2$ by $n(\zeta)$. Epstein [1986] defined the hyperbolic Gauss map by $G(\zeta) = n^+(\zeta)$, for $\zeta \in S$ There is a uniquely determined horosphere $\sigma(\zeta)$ based at $n(\zeta)$ that is tangent to S at ζ.

If we reverse the orientation of S we get another Gauss map $n^-(\zeta)$ that sends ζ to the "other" side of $\mathbb{S}^2$. When S is smoothly embedded and its principal curvatures satisfy $k_1, k_2 < 1$, the maps n^+, n^- are diffeomorphisms to disjoint open sets in $\mathbb{S}^2$. The composition $n^- \circ (n^+)^{-1}$ is a kind of "reflection" in $\partial S \subset \mathbb{S}^2$. The situation is studied in detail in [Epstein 1986].

1-39. *Fricke's Lemma* (see [Magnus 1980] for the history and a proof). Let A_1, A_2, A_3 be Möbius transformations. With Magnus introduce the notation

$$x_\nu = \operatorname{tr} A_\nu, \quad y_{\nu\mu} = \operatorname{tr} A_\nu A_\mu, \quad z_{\nu\mu\sigma} = \operatorname{tr} A_\nu A_\mu A_\sigma,$$

$$P = x_1 y_{23} + x_2 y_{13} + x_3 y_{12} - x_1 x_2 x_3,$$

$$Q = x_1^2 + x_2^2 + x_3^2 + y_{12}^2 + y_{13}^2 + y_{23}^2 + y_{12} y_{13} y_{23} - x_1 x_2 y_{12} - x_1 x_3 y_{13} - x_2 x_3 y_{23} - 4.$$

Prove Fricke's Lemma, namely the formula

$$P = z_{123} + z_{132}, \quad Q = z_{123} z_{132}.$$

In other terms, z_{123} and z_{132} are the roots of the equation $z^2 - Pz + Q = 0$.

1-40. *Finer properties of isometric circles.* The following properties have proved very useful in Jørgensen's hands in analyzing one and two generator groups.

Lemma 1.6.2. *Let A and B be Möbius transformations on S^2 and let $\mathcal{I}$, $\mathcal{B}$ be as in Section 1.5.*

(i) $\mathcal{B}(B)$ *covers a set σ on the isometric plane or circle $\mathcal{I}(A)$ if and only if* $\mathcal{B}(BA^{-1})$ *covers $A(\sigma)$ on the isometric plane or circle $\mathcal{I}(A^{-1})$.*
(ii) *If the circle $\mathcal{I}(B)$ is internally tangent to $\mathcal{I}(A)$ at the point x, then $\mathcal{I}(BA^{-1})$ is externally tangent to $\mathcal{I}(A^{-1})$ at $A(x)$.*
(iii) $\mathcal{I}(A), \mathcal{I}(AB), \mathcal{I}(B)$ *have a common point x if and only if* $\mathcal{I}(A), \mathcal{I}(AB^{-1}), \mathcal{I}(B^{-1})$ *have a common point $B(x)$.*
(iv) *Suppose $\mathcal{I}(B_1), \dots, \mathcal{I}(B_n)$, $n \geq 3$, go through a point x, and that $\bigcup_1^n \mathcal{B}(B_i)$ covers a neighborhood of x. Then for each k, the circle $\mathcal{I}(B_k^{-1})$ and every circle $\mathcal{I}(B_i B_k^{-1})$, for $i \neq k$, pass through $B_k(x)$, and the union of their interiors covers a neighborhood of $B_k(x)$.*
(v) *The sum of the excesses* (see below) *of the three pairs of isometric circles (A, B), (A^{-1}, BA^{-1}) and (B^{-1}, AB^{-1}) is $12\pi - 2(\lambda_1 + \lambda_2 + \lambda_3)$, where $\lambda_1, \lambda_2, \lambda_3$ are the exterior angles of intersection of each of the three pairs of circles.*

If two circles bound overlapping open disks, the overlap is bounded by an arc of each one. The *excess* of the pair is defined as the sum of the (euclidean) central angles subtended at the center of the two circles by the *complements* of the arcs. If the circles do not intersect at all, the excess is defined as 4π.

Proof. Properties (1)–(4) follow from the chain rule,

$$(BA^{-1})'(Az) = \frac{B'(z)}{A'(z)}.$$

For $z \in \sigma$ we have $|A'(z)| = 1$ and $|B'(z)| > 1$, so $|(BA^{-1})'(Az)| > 1$. Conversely, if $|(BA^{-1})'(Az)| > 1$ and $|A'(z)| = 1$, then $A(z) \in \mathcal{I}(a)$ and necessarily $|B'(z)| > 1$. Assertions (3) and (4) are consequences and (2) is a limiting case.

For (5), the key step here is to remember a theorem from high school geometry. Consider a closed arc σ on a circle. Draw the rays from the center to its end points. Let 2θ denote the angle they subtend. Choose a point ζ on the circle, but not on σ and draw the lines from ζ to the end points of σ. Then the lines from ζ subtend the angle θ with σ. In the limiting case that one of the end points of σ approaches ζ, the angle θ approaches the angle between the remaining line from ζ and the tangent to the circle at ζ.

Next, suppose $\mathcal{B}(A) \cap \mathcal{B}(B) \neq \emptyset$. Let σ_A denote the arc $\mathcal{I}(A) \cap \mathcal{B}(B)$ and σ_B the arc $\mathcal{I}(B) \cap \mathcal{B}(A)$. Let θ_1 denote the angle subtended at the center $c(A)$ by the *complementary arc* $\mathcal{I}(A) \setminus \sigma_A$ and θ_2 subtended at $c(B)$ by $\mathcal{I}(B) \setminus \sigma_B$. The excess at this intersection is, by definition, $\theta_1 + \theta_2$.

Draw the straight line $\boldsymbol{l}$ through the two points $\mathcal{I}(A) \cap \mathcal{I}(B)$. For ease of reference assume $\boldsymbol{l}$ is a vertical line. Choose one of the points of intersection ζ and draw there the tangent lines to the two circles. Let λ_3 denote the *exterior* angle of intersection of the two circles, that is, the angle between the two tangents that lies exterior to both circles. We claim that the angle between $\boldsymbol{l}$ and the tangent line to $\mathcal{I}(A)$ is $\theta_1/2$ and correspondingly that between $\boldsymbol{l}$ and $\mathcal{I}(B)$ is $\theta_2/2$. This is a consequence of the limiting case of the high school theorem presented above.

Summing the angles at ζ shows that $\frac{1}{2}\theta_1 + \frac{1}{2}\theta_2 + \lambda_3 = 2\pi$.

Now pass on to the next pair of intersecting circles, $\mathcal{I}(A^{-1})$ and $\mathcal{I}(BA^{-1})$. Let $\sigma_{A^{-1}}$ and $\sigma_{BA^{-1}}$ denote the arcs $\mathcal{I}(A^{-1}) \cap \mathcal{B}(BA^{-1})$ and $\mathcal{I}(BA^{-1}) \cap \mathcal{B}(A^{-1})$. Recall that $\mathcal{I}(A)$ has the same radius as $\mathcal{I}(A^{-1})$ and that A is a euclidean isometry from the former to the latter. This means that the angle subtended at $c(A^{-1})$ by the complementary arc $\mathcal{I}(A^{-1}) \setminus \sigma_{A^{-1}}$ is again θ_1. Let θ_3 denote the angle subtended at $c(BA^{-1})$ by the complementary arc $\mathcal{I}(BA^{-1}) \setminus \sigma_{BA^{-1}}$. Denote the exterior angle of intersection of the two circles by λ_2. Again we find that $\frac{1}{2}\theta_1 + \frac{1}{2}\theta_3 + \lambda_2 = 2\pi$.

Once more, carry out this construction for the intersecting circles $\mathcal{I}(B^{-1})$ and $\mathcal{I}(AB^{-1})$. The angle subtended at $c(B^{-1})$ is now θ_2 while the angle at $c(AB^{-1})$ is θ_3. If λ_1 denotes the exterior angle of intersection of the circles, we find as before that $\frac{1}{2}\theta_2 + \frac{1}{2}\theta_3 + \lambda_1 = 2\pi$.

Putting the three calculations together we conclude that the *total excess* is

$$2(\theta_1 + \theta_2 + \theta_3) = 12\pi - 2(\lambda_1 + \lambda_2 + \lambda_3).$$

In the limiting case that the circles do not cross each other, the total excess is $3 \times 4\pi$. $\square$

2

Discrete groups

This chapter introduces the related notions of discreteness and discontinuity, limit set and ordinary set. We establish the connection between discrete groups of Möbius transformations and hyperbolic manifolds and orbifolds. Some classical special cases of discrete groups are presented: elementary groups (which we classify), fuchsian and Schottky groups. The chapter includes crash courses on covering manifolds, Riemann surfaces, and quasiconformal mappings. The first two of these topics help us understand the boundaries of the 3-manifolds, while the latter shows us how to make controlled deformations of them.

We start by recalling some notions from group theory. Two groups G, H of Möbius transformations are said to be *conjugate* if there is a Möbius transformation T such that $G = THT^{-1}$; in other words G is the group consisting of the elements ThT^{-1}, for $h \in H$. As we did with single Möbius transformations in Chapter 1, we will often find it convenient to "normalize" a group of transformations, replacing it by a representative of its conjugacy class for which we stipulate some propitious property.

If A, B are Möbius transformations, $\langle A, B\rangle$ denotes the group generated by A and B and $\langle A\rangle$ the cyclic group generated by A.

A group is *torsion-free* if no element apart from the identity has finite order. Thus a torsion-free group of Möbius transformations is one that has no elliptic elements.

If a group G acts on a set X, the *stabilizer* of a subset $\Sigma \subset X$ under G is the set

$$\mathrm{Stab}(\Sigma) = \mathrm{Stab}_G(\Sigma) = \{g \in G : g(\Sigma) = \Sigma\}.$$

The case that interests us is where G is a group of Möbius transformations and Σ is a subset of $\mathbb{S}^2$.

2.1 Convergence of Möbius transformations

Lemma 2.1.1 (Convergence of Möbius transformations). *Suppose $\{T_n\}$ is an infinite sequence of distinct Möbius transformations such that the corresponding fixed points p_n, q_n converge to $p, q \in \mathbb{S}^2$; here either $p_n = q_n$, or T_n is elliptic, or p_n is the repelling and q_n the attracting fixed point of T_n. There is a subsequence $\{T_k\}$ with one of the following properties.*

(i) *There exists a Möbius transformation T such that* $\lim T_k(z) = T(z)$ *uniformly on* $\mathbb{H}^3 \cup \mathbb{S}^2$ (*considered with the euclidean metric*), *or equivalently*, $T_k \to T$ *for suitable choices of the associated matrices.*

(ii) $\lim T_k(z) = q$ *for all* $z \neq p$, *uniformly on compact subsets of* $\mathbb{H}^3 \cup (\mathbb{S}^2 \setminus \{p\})$. *Also* $\lim T_k^{-1}(z) = p$ *for all* $z \neq q$, *uniformly on compact subsets of* $\mathbb{H}^3 \cup (\mathbb{S}^2 \setminus \{q\})$. *Possibly* $p = q$.

Examples: $\{z+n\}$, $\{k^n z\}$, $\{e^{i/n} z\}$, $\{a^2 z + n(1-a^2)\}$.

Before proving the lemma, we state as a corollary a stronger form of Montel's famous theorem on "normal families" (the original version requires three omitted values).

Corollary 2.1.2. *Suppose $\{T_n\}$ is an infinite sequence of distinct Möbius transformations and $U \subset \mathbb{S}^2$ is a connected open set. Suppose there are two distinct points ζ_1, ζ_2 in $\mathbb{S}^2$ such that $T_n(U)$ avoids ζ_1 and ζ_2, for all n. Then there is an infinite subsequence $\{T_m\}$ which converges on U, uniformly on compact subsets, to a Möbius transformation or to a constant.*

Proof of Lemma 2.1.1. Assume that $\{T_n\}$ is a sequence whose fixed points converge as described in Lemma 2.1.1, and assume it has no subsequence which converges to a Möbius transformation.

Case 1: $p \neq q$. Choose $\zeta \in \mathbb{C}$ distinct from p, q, p_n, q_n for all n. Set $R_n(z) = (z, \zeta, p_n, q_n)$ so that $\lim R_n(z) = R(z) = (z, \zeta, p, q)$, uniformly on $\mathbb{S}^2$.

The transformation $S_n(z) = R_n T_n R_n^{-1}(z)$ fixes $0, \infty$ and has the same convergence properties as $\{T_n\}$. We have for large indices $S_n(z) = a_n z$ with $|a_n| \geq 1$. If $|a_n|$ is bounded for infinitely many indices then a subsequence converges to a Möbius transformation. Otherwise there exists a subsequence $\{S_m\}$ for which $\lim a_m = \infty$. In this case, $\{S_m\}$ converges uniformly to ∞ outside any given neighborhood of $z = 0$.

Case 2: $p = q$. Choose $\zeta_1, \zeta_2 \neq q_n, q$ and $\zeta_1 \neq \zeta_2$. Set $R_n(z) = (z, \zeta_1, \zeta_2, q_n)$. Again $\lim R_n(z) = R(z) = (z, \zeta_1, \zeta_2, q)$. Set $S_n(z) = R_n T_n R_n^{-1}(z)$. This fixes ∞ and has the same convergence properties as $\{T_n\}$. So $S_n(z) = a_n z + b_n$; the other fixed point of S_n is $-b_n/(a_n - 1)$. If for a subsequence $\lim b_m = b \neq \infty$, then $\lim a_m = 1$. In this case $\lim S_m(z) = z + b$. If instead $\lim b_m = \infty$, rewrite S_m as

$$S_m(z) = b_m \left(\frac{(a_m - 1)z}{b_m} + 1 \right) + z.$$

Since $\lim (a_m - 1)/b_m = 0$, we have $\lim S_m(z) = \infty$ for all z. As for the inverse,

$$S_m^{-1}(z) = \frac{b_m}{a_m} \left(\frac{z}{b_m} - 1 \right).$$

Because

$$\lim \frac{a_m - 1}{b_m} = \lim \left(\frac{a_m}{b_m} - \frac{1}{b_m} \right) = 0$$

and $\lim b_m = \infty$, we find $\lim a_m/b_m = 0$. Therefore $\lim S_m^{-1}(z) = \infty$ as well, for all $z \in \mathbb{C}$. □

Proof of Corollary 2.1.2. Let $\{T_m\}$ be a convergent sequence as in Lemma 2.1.1. Suppose the limit is not a Möbius transformation. Then $\lim T_m(z) = q$ uniformly outside any given neighborhood of p.

Case 1: $p \neq q$. We may assume that $p = 0,\ q = \infty$. If $0 \notin U$, Corollary 2.1.2 is true. If $0 \in U$ then $p_m \in U$ for all large indices. We show that then, for all large indices, the equation $T_m(z) = 0$ has a solution in U. For choose a disk $D \subset U$ centered at 0. Given a smaller concentric disk D_0, the fixed points $\{p_m\}$ are contained in D_0 for all large indices, while the $\{q_m\}$ are in the exterior of D. The image disk $T_m(D)$ contains p_m but not q_m. And $\lim_{m\to\infty} T_m(\partial D) = \infty$. Therefore $T_m(D)$ covers 0, and in fact any given point $\zeta \in \mathbb{C}$, for all large indices.

In sum, if $p \in U$, then $p \neq \zeta_1, \zeta_2$. Furthermore, if $\zeta_i \neq \infty$, $T_m(D)$ covers $\{\zeta_i\}$ for all large indices, in contradiction to our assumption. We conclude that $p \notin U$, in which case there is a subsequence converging uniformly on compact subsets to q.

Case 2: p=q. We may assume that $p = q = \infty$. Choose a disk D centered at ∞ so that at least one of the points ζ_i does not lie in D. Since $\lim T_m^{-1}(z) = \infty$, uniformly in the complement of D, we have in particular $T_m^{-1}(\zeta_i) \in D$ for all large indices. Therefore $T_m(D)$ covers ζ_i for all large indices. Once again we conclude that $p = \infty \notin U$ so that $\{T_m\}$ converges uniformly to ∞ on compact subsets of U. □

The example of the powers of a loxodromic transformation acting on the complement U of its attracting fixed point shows that the hypotheses of Corollary 2.1.2 are best possible.

We also include in this section the following elementary fact.

Lemma 2.1.3. *If g is loxodromic and h exchanges the fixed points of g, then* $h^2 = \mathrm{id}$ $(\mathrm{tr}(h) = 0)$.

Proof. In any case h^2 fixes the fixed points of g. But h has its own fixed point or points which h^2 fixes as well. □

2.2 Discreteness

In this section we begin our study of groups of Möbius transformations. A group G of Möbius transformation is *discrete* if there is no infinite sequence of distinct elements in the group that converges to the identity. Using Lemma 2.1.1 we see that each of the following conditions is equivalent to discreteness.

(i) No infinite sequence of distinct elements of G converges to a Möbius transformation.
(ii) G acts *properly discontinuously* in $\mathbb{H}^3$: Given any closed ball $B \subset \mathbb{H}^3$, the set $\{g \in G : g(B) \cap B \neq \varnothing\}$ is finite.

(iii) G has no limit points in $\mathbb{H}^3$: Given $\vec{x} \in \mathbb{H}^3$, there is no point $\vec{y} \in \mathbb{H}^3$ with an infinite sequence of distinct elements $\{g_n\}$ in G such that $\lim g_n(\vec{y}) = \vec{x}$.

Proper discontinuity implies that discrete groups have at most a countable number of elements. To see this, exhaust $\mathbb{H}^3$ by a countably many closed balls $V_1 \subset V_2 \subset \cdots$ centered at some point $\mathcal{O} \in \mathbb{H}^3$. For each i, enumerate the at most finitely many elements $g \in G$ for which $g(O) \in V_i$.

A group G is called *elementary* if and only if it preserves one point or a pair of points on $\mathbb{S}^2$, or a point in $\mathbb{H}^3$. An equivalent definition is that a group is elementary if and only if any two elements of infinite order have a common fixed point; see Exercise 2-1.

It is difficult to determine whether the group generated by a given set of elements is discrete. An algorithm for deciding discreteness of two-generator groups in $\mathrm{PSL}(2, \mathbb{R})$ is presented in [Gilman 1995]. The best general result is the following necessary condition.

Jørgensen's Inequality [Jørgensen 1974b]. *If $G = \langle A, B \rangle$ is discrete then*

$$\left|\mathrm{tr}^2(A) - 4\right| + \left|\mathrm{tr}(ABA^{-1}B^{-1}) - 2\right| \geq 1, \tag{2.1}$$

except in the following three cases, which are elementary groups:

(i) G cyclic or a finite abelian extension of a cyclic group and $|\mathrm{tr}^2(A) - 4| < 1$.

(ii) A is loxodromic or elliptic with $|\mathrm{tr}^2(A) - 4| < \frac{1}{2}$ while B interchanges the fixed points of A.

(iii) A is parabolic while B is parabolic or elliptic of order 2, 3, 4 *or* 6 *and fixes the fixed point of A.*

Note that the left side of (2.1) depends continuously on the Möbius entries.

The inequality is often applied to show the impossibility of a situation that $\langle A_n, B_n \rangle$ remains nonelementary while $\lim A_n = \mathrm{id}$.

Jørgensen went on to draw the following conclusion [1977b].

Corollary 2.2.1. *A nonelementary group G is discrete if and only if every two-generator subgroup is discrete.*

If G preserves a disk in $\mathbb{S}^2$, then G is discrete if and only if every one-generator subgroup is discrete, that is, if and only if there are no elliptic transformations of infinite order.

In contrast, if a *nonelementary group H* is not discrete, Leon Greenberg [1962] proved that its closure in $\mathrm{PSL}(2, \mathbb{C})$, that is the set of all Möbius transformations which are limits of elements of H, is either the full group $\mathrm{PSL}(2, \mathbb{C})$ or it is the group of all Möbius transformations which preserve some round disk in $\mathbb{S}^2$.

Proof of Corollary 2.2.1. Assume G is nonelementary. We will see below (Corollary 4.1.5) that if all elements in G are elliptic, then G is elementary with a common fixed point in $\mathbb{H}^3$. Assuming this, there is a loxodromic or parabolic element B_1 in

G and an element C which neither shares a fixed point with B_1 nor interchanges the fixed points p, q of B_1, if loxodromic. Then the fixed points $C(p), C(q)$ of the loxodromic or parabolic $B_2 = CB_1C^{-1}$ are distinct from p, q. For all n we have $B_1^n(C(p)) \neq C(p)$ and $B_1^n(C(q)) \neq C(q)$, yet these points are as close as we please to one of p, q. Therefore for some n the fixed points of $B_3 = B_1^n B_2 B_1^{-n}$ are distinct from those of both B_1, B_2. That is, the three loxodromic or parabolic transformations B_1, B_2, B_3 have mutually distinct fixed points.

We claim that B_1, B_2, B_3 can be replaced if necessary by three other transformations with distinct fixed points, so that all of them are loxodromic. To confirm the claim, it is enough to show that B_1 and B_2 can be chosen not to be parabolic. This is a consequence of the following argument.

If X, Y are parabolic without a common fixed point we may assume $X = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $Y = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in normalized matrices. Then $\mathrm{tr}(X^n Y) = a + d + nc$, so that $Z = X^n Y$ will be loxodromic for all large n and will not share a fixed point with Y. Consequently YZY^{-1} is loxodromic as well and does not share a fixed point with Z.

Now assume that G is not discrete. We will show the existence of a two-generator subgroup that is not discrete. This will show that if every two-generator subgroup of a group is discrete, the group itself must be discrete.

So assume that there is an infinite sequence $\{A_n\}$ of distinct elements of G with $\lim A_n = \mathrm{id}$. For n sufficiently large,

$$\left|\mathrm{tr}^2(A_n) - 4\right| + \left|\mathrm{tr}(A_n B_i A_n^{-1} B_i^{-1}) - 2\right| < 1,$$

for $i = 1, 2, 3$. For each n, at least one element of B_1, B_2, B_3 does not share a fixed point with A_n. Passing to a subsequence if necessary we may assume say B_1 does not share a fixed point with any A_n. The group $\langle A_n, B_1\rangle$ is not elementary provided that A_n does not exchange the fixed points of B_1. Since $\lim A_n = \mathrm{id}$, the order of A_n, if finite must increase to ∞. For large enough n, A_n, B_1 do not satisfy Jørgensen's inequality; therefore $\langle A_n, B_1\rangle$ cannot be discrete.

Now assume that G is a nonelementary group preserving the unit disk $\mathbb{D}$. Suppose G is not discrete. We claim that G then contains an elliptic element of infinite order.

By the previous result we may assume that G is a two-generator group. By Selberg's lemma below (page 68), there is a finitely generated subgroup G_0 of finite index without elliptic transformations of finite order. This too is nonelementary and nondiscrete. Now we call on a theorem of C. L. Siegel repeated in [Lehner 1964, III.3J] and that we will ask the reader to prove in Exercise 2-3, that establishes that G_0 must contain elliptic elements of arbitrarily high order. In fact then, G_0 must contain elliptic elements of infinite order. □

Proof of Jørgensen's inequality. We will follow the original proof. Assume the inequality fails to hold, so that for $A, B \neq \mathrm{id}$ generating a discrete group,

$$\mu = \left|\mathrm{tr}^2(A) - 4\right| + \left|\mathrm{tr}(ABA^{-1}B^{-1}) - 2\right| < 1.$$

We study the sequence obtained by setting $T_0 = B$ and define inductively

$$T_n = T_{n-1} A T_{n-1}{}^{-1} = \begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix}, \quad a_n d_n - b_n c_n = 1.$$

Case 1: $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ is parabolic. Write

$$B = \begin{pmatrix} a_0 & b_0 \\ c_0 & d_0 \end{pmatrix},$$

where $a_0 d_0 - b_0 c_0 = 1$. We may assume B does not fix ∞ — otherwise $\langle A, B\rangle$ would be elementary and to be discrete, B would have to be parabolic or be elliptic of order 2, 3, 4 or 6 by Lemma 2.3.1(iii). Therefore $c_0 \neq 0$. Furthermore,

$$ABA^{-1}B^{-1} = \begin{pmatrix} 1 + a_0 c_0 + c_0{}^2 & 1 - a_0 c_0 - a_0{}^2 \\ c_0{}^2 & 1 - a_0 c_0 \end{pmatrix}, \quad \mathrm{tr}(ABA^{-1}B^{-1}) - 2 = c_0^2.$$

Therefore $\mu = |c_0^2| < 1$. We find for the sequence of conjugates $\{T_n\}$ that

$$\begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix} = \begin{pmatrix} 1 - a_{n-1} c_{n-1} & a_{n-1}{}^2 \\ -c_{n-1}{}^2 & 1 + a_{n-1} c_{n-1} \end{pmatrix}.$$

From this we deduce that

$$c_n = -c_0{}^{2^n}, \quad |c_n| = \mu^{2^{n-1}} < 1,$$

$$|a_{n+1} - 1| \leq |c_n|(n + |a_0|), \quad |d_{n+1} - 1| \leq |c_n|(n + |a_0|), \quad |b_{n+1} - 1| = |a_n{}^2 - 1|.$$

So $\lim c_n = 0$, $\lim a_n = \lim d_n = \lim b_n = 1$, and hence $\lim T_n = A$. Since $0 < |c_0| < 1$ we see from $c_n = -c_0{}^{2^n}$ that the elements of the sequence $\{c_n\}$, hence of the sequence of transformations $\{T_n\}$, are distinct. Therefore $\langle A, B\rangle$ is not discrete, a contradiction.

Case 2: $A = \left(\begin{smallmatrix} \rho & 0 \\ 0 & 1/\rho \end{smallmatrix}\right)$, with $|\rho| \geq 1$. Take B as before. We find that

$$\mathrm{tr}(ABA^{-1}B^{-1}) - 2 = -b_0 c_0 (\rho - \rho^{-1})^2, \quad (\rho - \rho^{-1})^2 = \mathrm{tr}^2(A) - 4.$$

Note in particular that $|\rho - \rho^{-1}|^2 \leq \mu < 1$.

Now

$$T_{n+1} = \begin{pmatrix} a_n d_n \rho - b_n c_n \rho^{-1} & a_n b_n (\rho^{-1} - \rho) \\ c_n d_n (\rho - \rho^{-1}) & a_n d_n \rho^{-1} - b_n c_n \rho \end{pmatrix}.$$

Consequently,

$$b_{n+1} c_{n+1} = -a_n b_n c_n d_n (\rho - \rho^{-1})^2 = -b_n c_n (1 + b_n c_n)(\rho - \rho^{-1})^2.$$

Inserting the formula for $\mathrm{tr}(ABA^{-1}B^{-1})$,

$$\begin{aligned} |b_1 c_1| &= |b_0 c_0|\, |(1 + b_0 c_0)(\rho - \rho^{-1})^2| \\ &= |b_0 c_0| \left| \mathrm{tr}^2(A) - 4 - \mathrm{tr}(ABA^{-1}B^{-1}) + 2 \right| \leq |b_0 c_0| \mu \leq |b_0 c_0|, \end{aligned}$$

since we are assuming $\mu < 1$. Using induction starting with the case $n = 1$ we just investigated, we find that $|b_n c_n| \leq |b_0 c_0| \mu^n$. Moreover the analysis shows that sequence $\{|b_n c_n|\}$ strictly decreases to 0, unless it equals zero after some point.

Note that $b_{n+1}/b_n = a_n(\rho^{-1} - \rho)$ and $c_{n+1}/c_n = d_n(\rho - \rho^{-1})$. Consequently if $b_{n+1} = c_{n+1} = 0$ while $b_n \neq 0$ and $c_n \neq 0$, necessarily $a_n = d_n = 0$ at $\mathrm{tr}(T_n) = 0$.

Case 2a: $b_n c_n \neq 0$, for all n. Since the sequence $\{b_n c_n\}$ is strictly decreasing, the elements $\{T_n\}$ are distinct. Because $\lim a_n d_n = 1$, from the formula for T_{n+1} we see that $\lim a_n = \rho$ and $\lim d_n = \rho^{-1}$. Again using the formula for T_{n+1} we find that $\lim(b_{n+1}/b_n) = \rho(\rho^{-1} - \rho)$ and $\lim c_{n+1}/c_n = \rho^{-1}(\rho - \rho^{-1})$.

If A is elliptic, that is if $|\rho| = 1$, the ratios $|b_{n+1}|/|b_n|$ and $|c_{n+1}|/|c_n|$ are approximately $|\rho - \rho^{-1}| < 1$ and therefore $\lim b_n = \lim c_n = 0$. Consequently $\lim T_n = A$, contradicting discreteness.

Consider more generally the transformations

$$S_n = A^{-n} T_{2n} A^n = \begin{pmatrix} a_{2n} & \rho^{-2n} b_{2n} \\ \rho^{2n} c_{2n} & d_{2n} \end{pmatrix}.$$

Again from the formula for T_{n+1}, the ratios $|b_{2n}|/|b_{2n-2}|$ and $|c_{2n}|/|c_{2n-2}|$ are approximately $|\rho - \rho^{-1}|^2 < 1$. Therefore $\lim S_n = A$, again a contradiction to discreteness.

Case 2b: $b_n c_n = 0$, $n \geq N$. For $n \geq N$, A and T_n share a fixed point.

If A is elliptic, its order exceeds 6. This is because $\mu < 1$ implies $\sin\theta < \frac{1}{2}$ since since $\mathrm{tr}^2(A) - 4 = 4\sin^2\theta$ and A has the form $z \mapsto e^{2i\theta} z$. If A and T_n share exactly one fixed point, then by Lemma 2.3.1 G is not discrete, a contradiction.

If A is loxodromic and shares exactly one fixed point with T_n then $\langle A, T_n\rangle$ cannot be discrete.

Therefore A and T_n, $n \geq N$, have the same pair of fixed points $0, \infty$.

If $N = 0$, $G = \langle A, B\rangle$ is a discrete elementary group.

If $N = 1$, then $a_0 = d_0 = 0$ and G is elementary.

Suppose $N \geq 2$ so that T_{N-1} is conjugate to A. Then $\mathrm{tr}(T_{N-1}) = \mathrm{tr}(A) = 0$. But then $\mu \geq 4$, contrary to our assumption.

Further analysis yields the itemization of elementary groups for Jørgensen's inequality. □

In particular the group $\left\langle \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\right\rangle$, where $ad - bc = 1$, is not discrete when $0 < |c| < 1$.

2.3 Elementary discrete groups

A loxodromic or elliptic element g in a discrete group G is called *primitive* if g is a generator of the cyclic subgroup consisting of all loxodromic or elliptic elements in G having the same fixed points (and axis) as g.

The purpose of this section is to present the classical classification [Ford 1929] of the elementary discrete groups. We will see that a discrete group is elementary if and only if it is either finite, abelian or it contains an abelian subgroup of finite index.

Finite groups

If G is a finite group, it consists only of elliptic transformations and there is a common fixed point in $\mathbb{H}^3$ (Corollary 4.1.8). If the group is not cyclic there is exactly one fixed point. The common fixed point may be taken as the origin of the ball model so that G becomes a group of rotations of $\mathbb{S}^2$. Thus G is the group of orientation preserving symmetries of a regular figure inscribed in $\mathbb{S}^2$: one of the platonic solids, or else an equatorial regular polygon. More specifically, we have the following cases: the *tetrahedral group* of order 12, which preserves (collectively) the set of vertices of a tetrahedron; the *octahedral group* of order 24, which preserves the vertices of an octahedron, or those of its dual cube; the *icosahedral group* of order 60, which preserves the vertices of an icosahedron or dodecahedron; and the *dihedral group* of order $2n$, for $n \geq 2$, which preserves the dihedron, the degenerate "solid" consisting of two coincident faces in the shape of an n-sided regular polygon inscribed in the equator. The dihedral group contains, in addition to rotations by $2\pi i/n$ about the center, rotations of order two about any diameter from a vertex or the midpoint of a side. Each of these groups is generated by two elements.

For any finite G, the sphere $\mathbb{S}^2$ is a branched covering of $\mathbb{S}^2/G \cong \mathbb{S}^2$ with branching orders $r_i \geq 2$. If G is cyclic, there are two branch points, the fixed points of G. If G is not cyclic, the branching is over three points having the following orders (r_1, r_2, r_3): $(2, 3, 3)$ for the tetrahedral group, $(2, 3, 4)$ for the octahedral group, $(2, 3, 5)$ for the icosahedral group, and $(2, 2, n)$ for the dihedral group. See 3-1, 2-26.

More generally, we will show in Lemma 4.1.5 that if H is an arbitrary group consisting entirely of elliptic transformations, then H is conjugate to a group of rotations of $\mathbb{S}^2$.

Infinite elementary discrete groups

An elementary discrete group G that is not finite has one of two additional properties: (1) G fixes a single point ζ on $\mathbb{S}^2$; or (2) G fixes a pair of points on $\mathbb{S}^2$.

(1) *One fixed point.* Here G contains only parabolic transformations and elliptic transformations all sharing the same fixed point, say ∞. The parabolic subgroup G_0 is either cyclic and conjugate to $\langle z \mapsto z+1 \rangle$, or it is a free abelian group of rank two and conjugate to $\langle z \mapsto z+1,\ z \mapsto z+\tau \rangle$, for some $\tau \in \mathbb{C}$ with $\operatorname{Im}\tau > 0$. See Exercise 2-4.

In the cyclic case G itself can be the finite extension by an elliptic of order two.

In the rank-two case, G is a finite extension of G_0 by elliptics fixing ∞, of order not exceeding six by Lemma 2.3.1(iii). The possibilities are $(2, 2, 2, 2)$ and $(3, 3, 3)$, $(2, 3, 6)$, $(2, 4, 4)$, meaning these are the orders of primitive elliptic elements, non-conjugate under G_0, which generate the four possible extensions. For each of the triples to arise, G_0 must have a special choice of τ. For details see [Ford 1929] and Exercise 2-27.

(2) *Two fixed points.* Here G is a finite extension of a cyclic loxodromic group with axis ℓ. It can be extended by an elliptic of finite order with rotation axis ℓ and extended

once again by an elliptic of order two which exchanges the endpoints. All these groups preserve ℓ.

Lemma 2.3.1. *Let G be an infinite group of Möbius transformations.*

(*i*) *If G is discrete, G is elementary if and only if it is a finite extension of an abelian group.*
(*ii*) *If $g \in G$ is loxodromic and $h \in G$ has exactly one fixed point in common with g then G is not discrete.*
(*iii*) *If $g_1 \in G$ is elliptic of order exceeding six and $g_2 \in G$ has exactly one fixed point in common with g_1, then G is not discrete.*
(*iv*) *If $g \neq \mathrm{id}$ is an element of a nonelementary discrete or nondiscrete group G, there is a loxodromic element in G without a common fixed point with g.*
(*v*) *A nonelementary discrete or nondiscrete group contains two loxodromic elements with no fixed points in common, and hence it contains infinitely many loxodromic elements with mutually distinct fixed points.*

Proof. Item (i) follows from the discussion above.

For (ii), we may assume that $g = \begin{pmatrix} \rho & 0 \\ 0 & 1/\rho \end{pmatrix}$ and $h = \begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix}$. We find that

$$g^n h g^{-n} h^{-1} = \begin{pmatrix} 1 & -ab(1-\rho^{2n}) \\ 0 & 1 \end{pmatrix}.$$

If $|\rho| < 1$ let $n \to +\infty$. If $|\rho| > 1$, let $n \to -\infty$. In either case, G cannot be discrete.

To prove (iii) suppose g_1 and g_2 have the common fixed point ∞. According to Lemma 1.5.2, their commutator $g_1 g_2 g_1^{-1} g_2^{-1}$ is parabolic, also with fixed point ∞. If $G = \langle g_1, g_2 \rangle$ is to be discrete, then the subgroup G_∞ of parabolic transformations fixing ∞ has a generator K whose period ω satisfies $|\omega| \le |\omega'|$ in comparison to the periods ω' of other elements of G_∞ (see [Ahlfors 1978]). Write $g_1(z) = az+b$, $|a| = 1$ and $K(z) = z+\omega$. Then $g_1 K g_1^{-1}(z) = z + a\omega$. In particular $|\omega| \le |a\omega - \omega|$ or $1 \le |a-1|$.

Now $a = e^{i\theta}$, where $\theta = 2\pi k/m$ for some relatively prime $m, k \in \mathbb{Z}$, since if G is to be discrete the elliptic elements have finite order. We may choose g_1 so that $\theta = 2\pi/m > 0$ and then $|a-1| = 2\sin(\pi/m)$. If $|a-1|$ is to be ≥ 1, then we must have $m \le 6$, where $m = 6$ gives equality.

The proof of (iv) involves three cases. We will show later in Corollary 4.1.5 that a nonelementary group, discrete or not, contains nonelliptic elements.

Case 1. $g = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is parabolic. There exists $h \in G$ without a common fixed point with g: $h = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $c \neq 0$. We find that $\mathrm{tr}(g^n h) = (a+d) + nc$. Thus for all large $|n|$, $g^n h$ is loxodromic and does not share a fixed point with g.

Case 2. $g = \begin{pmatrix} \rho & 0 \\ 0 & \rho^{-1} \end{pmatrix}$ is loxodromic, $|\rho| > 1$. We have to show there is an element $h \in G$ which does not share one of the fixed points p, q of g. Not all elements of G can fix say p, but perhaps there is one h_p which fixes only p and another h_q that fixes only q. But then $h = h_q h_p$ fixes neither. In addition h does not exchange p, q.

Since $\mathrm{tr}(g^n h) = a\rho^n + d\rho^{-n}$, $g^n h$ is loxodromic for most n, and does not share a fixed point with g.

Case 3. $g = \begin{pmatrix} \rho & 0 \\ 0 & \rho^{-1} \end{pmatrix}$, $|\rho| = 1$, is elliptic with fixed points $p = 0, q = \infty$. If there is a loxodromic $h \in G$ which does not share a fixed point with g we are done. If there is a parabolic $h \in G$ which does not share a fixed point, then $h^n g$ is loxodromic for all large $|n|$. Moreover it does not share a fixed point with g. Finally if $h \in G$ shares exactly one fixed point with g then either h is parabolic or $ghg^{-1}h^{-1}$ is parabolic. So assume h_p is parabolic and fixes p while h_q is parabolic and fixes p. Then for all large $|n|$, $h_p^n h_q$ is loxodromic and fixes neither.

Item (v) follows from (iv). For given $g \neq \mathrm{id}$ let $h \in G$ be loxodromic without a common fixed point with g. Then the fixed points of the loxodromic element ghg^{-1} are $g(p), g(q)$ where p, q are the fixed points of h. Unless g is elliptic of order two and exchanges p, q, the fixed points of ghg^{-1} will be distinct. If g exchanges the fixed points of h the subgroup $\langle g, h\rangle$ is elementary. Yet there is some element $g_1 \in G$ which does not fix or exchange the fixed points of h. Now we can use $g_1 h g_1^{-1}$. Once we have two, we can keep conjugating so as to get infinitely many. □

2.4 Kleinian groups

Discrete groups of Möbius transformations are called *kleinian groups*. To avoid special cases, a kleinian group is often assumed to be nonelementary as well. A kleinian group that preserves the interior (hence also the exterior) of a round disk on $\mathbb{S}^2$ is called a *fuchsian group*. Typically, a fuchsian group is taken to act on the unit disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ or on the upper half-plane $\{z \in \mathbb{C} : \operatorname{Im} z > 0\}$.

We know that a group is discrete if and only if it is properly discontinuous on $\mathbb{H}^3$. Therefore we focus our attention on $\mathbb{S}^2$ and make the following definition.

A point $\zeta \in \mathbb{S}^2$ is a *limit point* of the discrete group G if there exists $\xi \in \mathbb{S}^2$ such that $\lim T_n(\xi) = \zeta$, for an infinite sequence of distinct elements $\{T_n\} \in G$. The set

$$\Lambda(G) = \{\zeta \in \mathbb{S}^2 : \zeta \text{ is a limit point}\}$$

is called the *limit set* of G. It contains all loxodromic and parabolic fixed points. It is automatically invariant under G. If $\Lambda(G)$ contains no, one, or two points, G is an elementary group.

Lemma 2.4.1 (Properties of the limit set). *Suppose G is nonelementary, so $\Lambda(G)$ contains at least three points.*

(*i*) *The G-orbit of any $\zeta \in \Lambda(G)$ is dense in $\Lambda(G)$.*
(*ii*) *$\Lambda(G)$ is the closure of the set of loxodromic fixed points, and if there are parabolics, it is the closure of the set of parabolic fixed points as well.*
(*iii*) *$\Lambda(G)$ is a closed set.*
(*iv*) *The G-orbit of any point $x \in \mathbb{H}^3 \cup \mathbb{S}^2$ accumulates onto $\Lambda(G)$.*
(*v*) *If $D_1, D_2 \in \mathbb{S}^2$ are two open disks with disjoint closures, each of which meets $\Lambda(G)$, there exists a loxodromic element in G with a fixed point in D_1 and in D_2.*
(*vi*) *$\Lambda(G)$ is a perfect set (it has no isolated points).*

(vii) *Either* $\Lambda(G) = \mathbb{S}^2$ *or its interior is empty.*

(viii) *If* G_0 *has finite index in* G, *or if* G_0 *is a normal subgroup of* G, *then* $\Lambda(G_0) = \Lambda(G)$.

Proof. Suppose $\zeta \in \Lambda(G)$ and let D be an open disk centered at ζ. For some $w \in \mathbb{S}^2$ we have $\zeta = \lim A_n(w)$, where $\{A_n\} \in G$ is an infinite sequence of distinct elements.

Choose $\zeta_1 \neq \zeta_2$ in $\Lambda(G)$. We claim that the family $\{G\}$ of Möbius transformations acting in D cannot omit the two values ζ_1, ζ_2. Assume otherwise. Then by Montel's Theorem (Corollary 2.1.2), $\{G\}$ is a normal family on D.

However, for some index N, we have $A_n(w) \in D$ for all $n \geq N$. Set $B_n = A_n A_N^{-1}$ and $w' = A_N(w)$. Since $\lim B_n(w') = \zeta$ and we must have convergence, $\lim B_n(z) = \zeta$, uniformly on compact subsets of D. In particular, given a subdisk $\zeta \in D' \subset \overline{D'} \subset D$, for all large indices $B_n(D')$ is a proper subset of D'. Therefore for each large index n and $k > 1$, we have $B_n^k(D') \subset B_n^{k-1}(D') \subset \cdots \subset D'$. This can only happen for a loxodromic transformation with attracting fixed point in D'. Thus, for all large indices, B_n is loxodromic and ζ must be the limit of the attracting fixed points. But then, for a fixed large n, the sequence $\{B_n^{-k}\}_{k=1}^{\infty}$ does not converge uniformly on compact subsets of D because it contains the repelling fixed points. We have found a contradiction.

Suppose that, for some $\xi \in \Lambda(G)$, ζ is not a limit point of the G-orbit $G(\xi)$. Then there is a disk D centered at ζ that contains no point of $G(\xi)$; in other terms, the G-orbit of D does not meet ξ nor any other point of its orbit. We have just shown that this is impossible. This argument proves (i). It also proves (vi).

Since G is nonelementary, there are infinitely many distinct loxodromic transformations in G. If ξ is a fixed point of the loxodromic T, any point $A(\xi)$ in its G-orbit is a fixed point of a loxodromic ATA^{-1}. The same holds if ξ is a parabolic fixed point.

In addition, the closure of the set of loxodromic fixed points lies in $\Lambda(G)$. Indeed, letting q_n be the attracting fixed point of the loxodromic $T_n \in G$, we see that the limit $\zeta = \lim q_n$ is the limit of a subsequence of the set of positive powers $\{T_n^k(w)\}$ for any $w \in \mathbb{C}$ distinct from $p = \lim p_n$. Therefore (ii) and (iii) hold. This argument also shows that property (iv) holds.

If $\Lambda(G)$ is not all of $\mathbb{S}^2$ there is an open set U in its complement. Every loxodromic fixed point is a limit point of the G-orbit of U, and then so is every point of $\Lambda(G)$. Therefore $\Lambda(G)$ can have no interior, as was claimed in (vii).

To prove (v) (after [Beardon 1983, Theorem 5.3.8]), choose loxodromics $A_1 \in G$ with attracting fixed point in D_1 and $A_2 \in G$ with attracting fixed point in D_2. If the repelling fixed point of A_2 is in D_1 we are finished, so assume that it is not. There is a loxodromic h with fixed points q_1, q_2 distinct from the fixed points of A_1, A_2. Its conjugate $B_1 = A_1^m h A_1^{-m}$, $m > 0$ has fixed points $A_1^m(q_1)$, $A_1^m(q_2)$. For sufficiently large m these will both lie in D_1; fix such an m.

Choose a closed disk $D_1' \subset D_1$ containing the repelling p but not the attracting fixed point q of B_1. Fix a large n such that A_2^n sends q into D_2.

We claim that we can choose $r > 0$ so large that $T = A_2^n B_1^r$ has the properties

$$T(\overline{D_2}) \subset D_2, \quad T^{-1}(\overline{D_1'}) \subset D_1'.$$

For take r large enough that $B_1^r(D_2)$ is so close to q that application of A_2^n then sends it properly into D_2. In the other direction $A_2^n(q) \notin D_1'$ says $q \notin A_2^{-n}(D_1')$ so we may increase r as needed so that $B_1^{-r} A_2^{-n}(D_1')$ is properly contained inside D_1', close to the repelling fixed point of B_1.

A transformation T with these mapping properties can only be loxodromic with attracting fixed point in D_2 and repelling in D_1'.

The proof of (viii) is as follows. If $A \in G$ is loxodromic and G_0 has finite index in G, then $A^k \in G_0$ for some $k > 0$. Therefore $\Lambda(G_0)$ has the same set of loxodromic fixed points as $\Lambda(G)$, so the closures of the sets are the same. Instead, suppose that G_0 is a normal subgroup of G so that $gG_0g^{-1} = G_0$ for all $g \in G$. Then G_0 cannot be an elementary subgroup. The g-image of the fixed points of $h \in G_0$ are the fixed points of $ghg^{-1} \in G_0$. Therefore $g\Lambda(G_0) = \Lambda(G_0)$ for all $g \in G$. Since $\Lambda(G_0) \subset \Lambda(G)$, the G-orbit of a fixed point of a loxodromic $h \in G_0$ is dense in $\Lambda(G)$; therefore the limit sets are identical. □

Each component of $\Lambda(G)$ which is not a circle or a point is a fractal set; see Exercises 2-14 and 3-20.

The complementary open set,

$$\Omega(G) = \mathbb{S}^2 \setminus \Lambda(G),$$

is called the *ordinary set* or *regular set*, or *set of discontinuity*. Like $\Lambda(G)$, $\Omega(G)$ is preserved by G. It is the largest open subset of $\mathbb{S}^2$ on which G acts properly discontinuously.

Lemma 2.4.2 (Properties of the ordinary set). *Assume that G is finitely generated and not elementary, and that $\Omega(G) \neq \varnothing$.*

(*i*) *$\Omega(G)$ has one, two, or infinitely many components.*
(*ii*) *Each component of $\Omega(G)$ is either simply or infinitely connected.*
(*iii*) *If each of two components Ω_1, Ω_2 of $\Omega(G)$ is preserved by G, then each one is simply connected and $\Omega(G) = \Omega_1 \cup \Omega_2$.*
(*iv*) *If one component Ω of $\Omega(G)$ is preserved by G, all the others are simply connected.*

Proof. To prove (ii), assume a component Ω is finitely but not simply connected. At this point we have to anticipate the Ahlfors Finiteness Theorem (page 105) to assert that Ω is preserved by an element $g \in G$ of infinite order (such a g may not exist if G is not finitely generated). Choose a simple loop $\sigma \subset \Omega$ that separates the boundary components. The simple loops $\{g^k(\sigma)\} \subset \Omega$ converge to the fixed points or point of g. But each simple loop $g^k(\sigma)$ separates boundary components of Ω. Hence the fixed points are limits of infinitely many boundary components of Ω, a contradiction.

To prove (i), suppose there are a finite number of components $\Omega_1, \ldots, \Omega_m$ of $\Omega(G)$; we may assume that $\infty \in \Omega_m$. There is a subgroup G_0 of finite index and with the same limit set that preserves each of them.

Choose a loxodromic transformation $g \in G_0$. Since g in particular preserves Ω_1 and Ω_2, we can find simple arcs $\sigma_i \in \Omega_i$, $i = 1, 2$, such that $\sigma_i^* = \bigcup_{k=-\infty}^{\infty} g^k(\sigma_i)$ forms a simple arc in Ω_i between the two fixed points of g. This is most easily done by using the quotient surface $\Omega_i/\mathrm{Stab}(\Omega_i)$. Then $\sigma^* = \sigma_1^* \cup \sigma_2^* \cup \{p, q\}$ forms a simple closed curve meeting $\Omega(G_0)$ only in Ω_1 and Ω_2. Consider the two components of $\mathbb{S}^2 \setminus \sigma^*$. One of them, say U, contains Ω_m and ∞. The other, U', contains points of $\Lambda(G_0)$, for otherwise σ_1^* and σ_2^* could be connected by an arc that does not meet $\Lambda(G_0)$. Therefore we can find a loxodromic element $h \in G_0$ with attracting fixed point in U'. Connect ∞ to $h(\infty)$ by an arc $\tau \subset \Omega_m$ and set $\tau^* = \bigcup_{k=0}^{+\infty} h^k(\tau)$. Now τ^* is an arc in Ω_m connecting $\infty \in U$ to the attracting fixed point of h in U', so τ^* must cross σ^*, giving a contradiction.

Item (iii) also depends on Ahlfors' theorem. Using that the simplest proof involves 3-dimensional topology. We will present it in Section 3.8.

Item (iv) is a consequence of the fact that $\Lambda(G) = \partial\Omega$. The analysis in terms of three-dimensional topology is suggested in Exercise 3-11. □

It is relevant to refer again to L. Greenberg's theorem [1962], which has the following consequence. Suppose $\Omega \neq \mathbb{S}^2$ is a connected open set which is not a round disk. Then the group of all Möbius transformations which map Ω onto itself is either discrete or elementary, as in the case of a horizontal strip. Usually it will consist only of the identity.

The term *function group* is usually reserved for a group G with the property that $\Omega(G)$ has an infinitely connected component Ω that is invariant under G. The term arises from the fact that functions invariant under G can be constructed on Ω. The finitely generated function groups can be completely classified [Maskit 1988] or by topology; see [Marden 1977] and Exercise 3-11.

When G is not finitely generated, if two of the components of $\Omega(G)$ are invariant under G then as before they are both simply connected. Yet there may also be other components; each of these is also simply connected, but its stabilizer consists only of the identity (see [Accola 1966] or apply 3-dimensional topology as in Section 3.8). An example of Accola shows that indeed there can be infinitely many other components, which he called “atoms”. (The situation is reminiscent of the classical construction in point set topology known as the *lakes of Wada*: a family of three — or any number up to countably infinite — simply connected open sets on $\mathbb{S}^2$ each of which has the same boundary, namely the complement in $\mathbb{S}^2$ of the union of the open sets. See [Hocking and Young 1961, pp. 143–145].) However, when G is finitely generated, atoms cannot occur, as we will see from the Ahlfors Finiteness Theorem.

Here is an answer to the question of the $\pm$ ambiguity as we go from a group of Möbius transformations to a set of associated matrices in $\mathrm{SL}(2,\mathbb{C})$. In short, for discrete groups, the signs can be chosen unambiguously, except if there are elements of order two.

Theorem 2.4.3 [Culler 1986]. *A discrete group G can be lifted to an isomorphic group of matrices in* $\mathrm{SL}(2,\mathbb{C})$ *if and only if G has no elements of order two.*

This result is the best that can be hoped for, since the matrices corresponding to Möbius transformations of order two have order four — for example the normalized matrix that corresponds to $z \mapsto 1/z$ has order four. On the other hand, one can ask, with John Fay, whether any group can be lifted to an unnormalized matrix group in $\mathrm{GL}(2,\mathbb{C})$. (For example, the unnormalized matrix $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ does have order two.) The answer is not known, to my knowledge.

2.5 Quotient manifolds and orbifolds

A kleinian group G is usually best studied by studying its quotient space:

$$\mathcal{M}(G) = \mathbb{H}^3 \cup \Omega(G)/G, \quad \partial\mathcal{M}(G) = \Omega(G)/G,$$

namely, the set of equivalence classes

$$\{\{x\} : x \in \mathbb{H}^3 \cup \Omega \text{ with } x \equiv x_1 \text{ if and only } x_1 = g(x),\, g \in G\}.$$

The projection $x \to \{x\}$ is denoted by π.

We will often switch between thinking of a situation in $\mathbb{H}^3 \cup \Omega(G)$ and thinking of it in the quotient $\pi : \mathbb{H}^3 \cup \Omega \to \mathcal{M}$.

If G is torsion-free (no elliptics), then $\mathcal{M}(G)$ is an oriented* manifold with boundary $\partial\mathcal{M}(G)$, which may be empty. The projection π is a local homeomorphism $\mathbb{H}^3 \to \mathbb{H}^3/G$ and $\Omega \to \partial\mathcal{M}(G)$, because of proper discontinuity of the group action. The interior $\mathcal{M}(G)^{int} = \mathbb{H}^3/G$ has a complete hyperbolic structure arising from the projection of the hyperbolic metric in $\mathbb{H}^3$. Its fundamental group $\pi_1(\mathcal{M})$ is isomorphic to G. If we lived inside $\mathcal{M}^{int}$, then we would *see* the universal cover $\mathbb{H}^3$ as it is the space of light rays (geodesic rays) of $\mathcal{M}^{int}$ meeting our eye, since light follows the shortest paths. What we would see standing at a point $\pi(x) \in \mathcal{M}^{int}$ is the picture at $x \in \mathbb{H}^3$. This is strikingly demonstrated in the video *Not Knot* [Gunn and Maxwell 1991].

The name "hyperbolic manifold" is reserved for those $\mathcal{M}(G)$ *arising from groups G without elliptics.*

On the other hand if G contains elliptics, $\mathcal{M}(G)$ is called an *orbifold*. The additional structure of orbifolds will be described below.

$*$ For the record we point out that if G were a group with orientation reversing elements, the subgroup of orientation preserving elements would form a normal subgroup of index two. The corresponding nonorientable quotient manifold would have a two-sheeted cover which is orientable and an orientation reversing isometry which interchanges the sheets.

Manifolds and their coverings

We will briefly review salient aspects of the theory of coverings of oriented surfaces and 3-manifolds thereby giving more insight to the nature of quotient spaces as well. Our applications will be to Riemann surfaces (Section 2.6) and hyperbolic manifolds, so our discussion will be carried out with these cases in mind. In particular the surfaces and manifolds will be oriented.

We will start by focusing on surfaces.*

Associated with a surface S and a given basepoint $O \in S$ is the *fundamental group* $\pi_1(S; O)$ of homotopy classes of closed paths from O. Choose a subgroup H of $G = \pi_1(S; O)$. For example, H may be the cyclic group generated by a single loop, or it may be the identity. A more interesting example is the commutator subgroup, which is the subgroup generated by commutators of pairs of elements of G (the subgroup is also called the homology group since it corresponds to the elements in $\pi_1(S; O)$ which are homologous to zero).

Corresponding to H is the regular† covering surface S_H constructed as follows. Consider equivalence classes of pairs $\{(z; \alpha_z)\}$, where $z \in S$ and α_z is a path from O to z. The pairs associated with paths α_1, α_2 from O to z are equivalent if the homotopy class of $\alpha_2^{-1}\alpha_1$ is in H. In particular $(z, \alpha_z) \equiv (z, \alpha_z\gamma)$ if $\gamma \in H$. The surface S_H is the set of equivalence classes $\{(z, \alpha_z)\}$ with the topology determined from S as follows: A neighborhood N^* of (z, α_z) consists of the pairs $\{(w, \sigma_w\alpha_z)\}$, where w lies in a neighborhood N of z and σ_w is a path in N from z to w.

The map $\pi : (z, \alpha_z) \in S_H \mapsto z \in S$, called the projection, is a local homeomorphism of S_H onto S. The points in $\{\pi^{-1}(z)\}$ are said to *lie over* $z \in S$. If H has finite index‡ n in $G = \pi_1(S; O)$, S_H is n-sheeted over S—there are exactly n distinct points of S_H lying over each point of S.

The point $O^* \in S_H$ determined by the class (O, γ), $\gamma \sim 1$, is the corresponding basepoint of S_H; it (and many others) lies over O. *The fundamental group* $\pi_1(S_H; O^*)$ *is isomorphic to* H. If $H \neq$id is cyclic, so is the fundamental group of S_H; in this case S_H is homeomorphic to an annulus. If $H =$ id, then S_H is simply connected and is called the *universal covering surface* of S. If $H = G$ then $S_H = S$.

A map $f : S \to S$ that, say, fixes the basepoint O lifts to a map $S_H \to S_H$ if and only if f induces an automorphism of the subgroup H onto itself.

A *deck transformation* (also called a cover transformation) is a fixed point free, orientation preserving homeomorphism τ^* of S_H onto itself with the property that $\pi(\tau^*(x)) = \pi(x)$; that is for each point $z \in S$, τ^* interchanges the points lying over z. The group of deck transformations of $\mathcal{M}(H)$ over $\mathcal{M}(G)$ is isomorphic to the quotient group $N(H)/H$. Here $N(H) = \{g \in G : gHg^{-1} = H\}$ is called the *normalizer* of H in G. An element $\gamma \neq \mathrm{id} \in N(H)$ induces the cover transformation $(z, \alpha_z) \mapsto (z, \alpha_z\gamma)$.

* Formally a surface is a connected 2-dimensional manifold, that is a Hausdorff space with an open covering of sets homeomorphic to open sets in $\mathbb{C}$.

† A regular covering S^* is one with the property that if $\alpha \subset S$ is a closed arc, and $x^* \in S^*$ lies over its initial point, then α can be lifted in its entirety from x^*.

‡ H has index n in G if there are n distinct cosets $\{Hg_k\}$, $g_k \in G$, such that $G = \bigcup_k Hg_k$. In this case $\bigcap_{g\in G} gHg^{-1} = \bigcap_k g_k H g_k^{-1}$ is a normal subgroup of finite index in G.

If $N(H) = G$ then H is a called a *normal subgroup* of H and S_H is called a *normal covering*. In this case the group of deck transformations is isomorphic to G/H: Given any two points O_1^*, O_2^* over O, in a normal covering S_H there is a deck transformation taking $O_1^* \to O_2^*$. In particular, when $H = \{\text{id}\}$, the group of deck transformations of the universal cover is isomorphic to the fundamental group G. Another normal covering is generated by the commutator subgroup. In this case the group of deck transformations is isomorphic to G/H and to the first homology group of S, this is a free abelian group of rank $2g$ if S is a closed surface.

In general however, there may or may not be deck transformations; for example if H is cyclic and S is a closed surface of genus exceeding one, then $N(H) = H$ and there are none, yet S_H is infinite-sheeted over S.

We will have need of extending our definition to *branched covers* S^* of S. The difference here is that a *discrete* set of points $\{\zeta_i\} \subset S^*$ is distinguished. We then have regular coverings of the punctured surfaces $S^* \setminus \{\zeta_i\} \to S \setminus \{\pi(\zeta_i)\}$ on which π is a local homeomorphism. But in a small neighborhood N of each ζ_i, $\pi : S^* \to S$ is not a homeomorphism, rather it can be taken as the map $z \mapsto z^r$, where ζ_i corresponds to 0: In N the projection is r-to-1. The point $\zeta_i \in S^*$ is referred to as a *branch point* and its projection $\pi(\zeta_i)$ is the *branch value* or *cone point*. The integer $r = r(\zeta_i) \geq 1$ is the order of ramification ($r = 1$ stands for a regular point). Paths in S lift to S^* provided they avoid the cone points. That the branched cover has N sheets implies that if $\{x_i^*\} \subset S^*$ are the distinct points lying over $x \in S$, then

$$\sum r(x_i^*) = N.$$

In particular, if $N = 2$, there is at most one branch point over x; if there is one, its order is two.

The Euler characteristic of an oriented, compact, triangulated surface of genus $g \geq 0$ and $b \geq 0$ boundary components is

$$\chi(S) = V - E + T = 2 - 2g - b, \tag{2.2}$$

where V is the number of vertices, E the number of edges and T the number of triangles.

The precise relationship between the topologies of a surface and its covering is governed by the Riemann–Hurwitz formula. Suppose S^* is an N-sheeted cover of the compact surface S, where S has genus $g \geq 0$ and $b \geq 0$ boundary components. Triangulate S so that all the branch values are vertices, and assume there are no branch values on the boundary components. Lifting the triangles to S^*, we can compute $\chi(S^*)$ in terms of $\chi(S)$. The result is the *Riemann–Hurwitz formula*, $\chi(S^*) = N\chi(S) - \sum(r(x_i) - 1)$, where the sum is over all branch points on S^*. In a more useful form,

$$2g^* + b^* = 2(1 - N) + N(2g + b) + \sum(r(x_i) - 1). \tag{2.3}$$

Thus, if $S = \mathbb{S}^2$, the (finite) coverings are closed surfaces* satisfying $2g^* + 2N - 2 = \sum(r(x_i) - 1)$; a closed surface of any genus can be so constructed. For a torus, $g = 1, b = 0$, the corresponding formula is $g^* - 1 = \frac{1}{2}\sum(r(x_i) - 1)$. If each $r_i = 2$, the formula is $g^* - 1 + N = n/2$, where there are n branch points.

We now turn to the case of hyperbolic manifolds $\mathcal{M}(G)$. Covering manifolds (unbranched) correspond to subgroups H of G, that is, $\mathcal{M}(H)$ is a hyperbolic covering of $\mathcal{M}(G)$. The group of deck transformations is isomorphic to $N(H)/H$; the deck transformations are fixed point free, orientation preserving isometries. On the other hand, the group of orientation preserving isometries $\mathcal{M}(G) \to \mathcal{M}(G)$ is isomorphic to $N(G)/G$. Here $N(G)$ is the normalizer of G in the full group of all Möbius transformations. Now elements of $N(G)$ can be elliptic. However very often $\mathcal{M}(G)$ has no orientation preserving "symmetries", that is, $N(G) = G$.

Two hyperbolic manifolds $\mathcal{M}(G_1)$, $\mathcal{M}(G_2)$ are *commensurable* if $G_1 \cap G_2$ has finite index in both G_1 and G_2. In this case $\mathcal{M}(G_1 \cap G_2)$ is a finite-sheeted cover of both $\mathcal{M}(G_1)$ and $\mathcal{M}(G_2)$. Conversely if $\mathcal{M}(H)$ is a finite-sheeted cover of both $\mathcal{M}(G_1)$ and $\mathcal{M}(G_2)$ then H is conjugate to a subgroup of finite index in each of G_1 and G_2. The term can equally be applied to fuchsian and orbifold groups (see below).

The set of all Möbius transformations g which have the property that gGg^{-1} is commensurable with G, is called the *commensurator* $C(G)$ of G. It is a group as well since if $g_1gGg^{-1}g_1^{-1}$ has finite index in gGg^{-1} and gGg^{-1} has finite index in G then $g_1gGg^{-1}g_1^{-1}$ also has finite index in G.

Contrast the commensurator $C(G)$ of G with its normalizer $N(G)$ which is a subgroup of $C(G)$. The normalizer contains all (orientation preserving) isometries of $\mathcal{M}(G)$. The commensurator consists of all orientation preserving isometries of all finite-sheeted covers of $\mathcal{M}(G)$. For if $\mathcal{M}(H) \to \mathcal{M}(G)$ is a finite cover, then H is conjugate to a finite index subgroup of G; we may assume $H \subset G$. If the Möbius transformation T induces an automorphism of $\mathcal{M}(H)$, then $THT^{-1} = H$ so that $TGT^{-1} \cap G \supset H$. Since H also has finite index in TGT^{-1}, $TGT^{-1} \cap G$ has finite index in G. For more applications, see Exercise 3-14.

Orbifolds

Consider now the situation when G has elliptic elements. In this case the quotient $\mathcal{M}(G)$ is called an *orbifold*. The hyperbolic structure of $\mathcal{M}(G)$, which is also oriented, has mild singularities (see Exercise 2-2) along the projection of the totality of rotation axes of the elliptic elements. This projection is called the *singular set* or *branch locus* of the orbifold.

The projection of an elliptic axis ℓ is usually called a *cone axis* as it is reminiscent of the paper-and-scissors construction of a cone by wrapping up a wedge of angle $< 2\pi$; correspondingly the points on the cone axis are called *cone points*. Locally the projection has the form $(z, t) \mapsto (z^r, t)$, where t is a coordinate along the rotation axis and z is a complex coordinate in a plane orthogonal to the axis. The *cone angle*

* A closed surface or manifold is one which is compact, without boundary.

$2\pi/r$ assigned to $\pi(\ell)$ is the angle of rotation of a primitive element—a generator of the cyclic subgroup that has rotation axis ℓ. The Isolation of Cone Axes property in Theorem 3.3.4 gives additional information about the separation of cone axes in $\mathcal{M}(G)$.

An elliptic rotation axis γ^* may also be the axis of a loxodromic in G. In this case ℓ will project to a simple loop γ in the singular set of the quotient. There may also be other elliptic axes that intersect γ^*. If so, the common point of intersection is stabilized by a finite elliptic group.

To better understand the structure of the singular set, we will start with the case of a finite group G.

Lemma 2.5.1. *For a finite group G associated with a regular solid, $\mathcal{M}(G)$ is topologically a closed ball, and there exists a point $O \in \mathbb{H}^3/G$ from which exactly three cone axes emanate. Each cone axis from O ends at one of three branch points on $\partial\mathcal{M}(G) = \mathbb{S}^2/G$. The angles at O between the axes are uniquely determined by G.*

Conversely given three distinct points on $\mathbb{S}^2$ there are conjugates of G whose cone axes end at those points in any prescribed order.

An elliptic rotation axis either ends at a point of $\Omega(G)$, or at a parabolic fixed point $\zeta \in \partial\mathbb{H}^3$. In a discrete group, the subgroup of parabolics that fix ζ is either cyclic, or it is free abelian of rank two, as we found when we examined the elementary groups. Correspondingly, we will refer to ζ as a rank one or rank two parabolic fixed point. The conjugacy classes of parabolic fixed points give rise to certain structures in $\mathcal{M}(G)$ called cusps, to be described in detail in Section 3.2. Here it suffices to say that a geodesic ray in $\mathcal{M}(G)$ ends at a rank one or rank two cusp if and only if any lift to $\mathbb{H}^3$ ends at a rank one or rank two parabolic fixed point.

Here is a description of the singular set as a graph in the quotient orbifold:

Proposition 2.5.2. *In any kleinian $\mathcal{M}(H)$, the singular locus is a graph, and a component can be compact in Int($\mathcal{M}(G)$) or not. Each edge has an order $n \geq 2$. Emanating from each interior vertex are three edges of orders* $(2, 3, 3)$, $(2, 3, 4)$, $(2, 3, 5)$, *or* $(2, 2, n)$, *for $n \geq 2$.*

If an edge does not end at a vertex, it may end at a point on $\partial\mathcal{M}(G)$, or at a rank one or rank two cusp. If it ends at a rank one cusp, it must have order two. If it ends at a rank two cusp it may have order two; two or three additional edges can end there as well, in which case they have orders $(3, 3, 3)$, $(2, 3, 6)$, $(2, 4, 4)$, *or* $(2, 2, 2, 2)$.

Note that for the various cases, quotient of a small horoball or euclidean ball about the common fixed point results in a euclidean or spherical orbifold (compare with Section 3.2 and Exercise 3-1).

For an example, consider three mutually orthogonal lines ℓ_1, ℓ_2, ℓ_3 intersecting at a point, for example the three coordinate axes at the origin in the ball model. Consider the group G generated by 180° rotations about each line.

Next take a point $x \neq 0$ on say ℓ_1, and take an orthogonal system $\ell_1^*, \ell_2', \ell_3'$ through x. Let G' be the group generated by the 180° rotations about the three lines ℓ_1, ℓ_2', ℓ_3'.

Consider the group $H = \langle G, G' \rangle$. Now ℓ_1 is the axis of a loxodromic element T that maps 0 to $T(0)$ which have equal distance from x. The branch lines form a trivalent graph, three edges hit every vertex. A fundamental set consists of half-lines $\ell_2^+, \ell_3^+, \ell_{2'}^+, \ell_{3'}^+$ and, in addition the segment of ℓ_1 from 0 to x. Down in $\mathbb{H}^3/H$, the projection $\pi(\ell_1)$ is a line segment from $\pi(0)$ to $\pi(x)$, and back, a degenerate simple loop. From each of $\pi(0)$ and $\pi(x)$, there are two rays ending at points on the boundary.

The boundary of $\mathbb{H}^3/\langle T \rangle$ is a torus and on it are eight distinguished points, the endpoints of the projection of $\ell_2, \ell_3, \ell_2', \ell_3'$. There is an automorphism of order two of the torus, that has no fixed points, that takes four of these points to the other four. Also acting on the torus is a group of order four generated by two automorphisms of order two, each with four of the distinguished points as fixed points. The quotient of the torus with respect to this group of order four is the sphere, and the torus is a four-sheeted cover, branched over four points on the sphere. The boundary of the orbifold $\mathbb{H}^3/H$ is the sphere: The four branch values are the endpoints of the four singular loci $\pi(\ell_2), \pi(\ell_3), \pi(\ell_2'), \pi(\ell_3')$.

Proof of Proposition 2.5.2. The quotient of $\mathbb{S}^2$ under the groups of the regular solids is again $\mathbb{S}^2$ with exactly three branch values. These have orders (2, 3, 3) for the symmetries of a regular tetrahedron, (2, 3, 4) for the symmetries of a cube or octahedron, (2, 3, 5) for the symmetries of a icosahedron or dodecahedron, and $(2, 2, n)$ for a dihedral group. All of the groups are *triangle groups*, which have three generators each of finite order (Exercise 2-5). These statements follow from the formula of Exercise 3-1 with details given in Exercise 2-26.

As in the example above, a given rotation axis ℓ may be intersected by other rotation axes at succession of distinct points. Each intersection point is the fixed point of one of the standard finite groups.

The second statement also follows from our itemization of elementary groups. Applying Equation 2.3, we see that the torus is a two-sheeted cover of $\mathbb{S}^2$ in the case (2, 2, 2, 2), a three-sheeted cover in the case (3, 3, 3), or a four-sheeted cover in the cases (2, 4, 4) and (2, 3, 6). □

To distinguish a quotient with the extra structure of cone axes, Thurston coined the term *orbifold*. For kleinian groups with elliptics, $\mathbb{H}^3$ is a simply connected branched cover of the orbifold $\mathbb{H}^3/G$ and $\Omega(G)$ may or may not be branched over $\partial\mathcal{M}(G)$. Actually 3-orbifolds are manifolds too, but new local coordinates need to be introduced in neighborhoods of the singular edges and vertices that map them to euclidean balls.

We will reserve the term orbifold for the cases that a singular set — cone axes — exists. Some authors use it to include both manifolds and orbifolds. We have not considered the case of nonorientable orbifolds. Such an orbifold would result, for example, from a reflection in a plane in $\mathbb{H}^3$.

The conformal boundary

The "boundary" $\partial\mathcal{M}(G)$ is infinitely far away from any interior point in the hyperbolic metric on $\mathcal{M}(G)^{\text{int}}$, yet it is intimately related to the interior structure. The isometries and the geodesics extend to it. The infinitesimal 3-dimensional coordinate frame at each point in $\Omega(G)$, with one direction the interior normal to $\mathbb{S}^2$, projects to the corresponding frame in $\partial\mathcal{M}$, also with one direction the interior normal to $\partial\mathcal{M}$. The boundary $\partial\mathcal{M}(G)$ has a conformal structure induced from $\Omega(G) \subset \mathbb{S}^2$: it is a union of Riemann surfaces (see next section). For this reason it is often called the *conformal boundary* of $\mathcal{M}(G)$.

If G is not elementary, no component of $\partial\mathcal{M}$ can be a sphere or a torus. Tori (without cone points) are excluded because by the Uniformization Theorem (see next section), they can arise only if a component Ω of $\Omega(G)$ over the torus is Möbius equivalent to $\mathbb{C}$ if Ω is simply connected, or Möbius equivalent to $\mathbb{C} \setminus \{0\}$ if Ω is not.

2.5.1 Two fundamental algebraic theorems

Using the following purely algebraic fact, the quotient orbifolds can often be analyzed by analyzing manifolds. For every orbifold obtained from a finitely generated group has a finite-sheeted cover which is a manifold:

Selberg's Lemma [1960]. *Every finitely generated group of matrices in* $\mathrm{SL}(2, \mathbb{C})$ *has a finitely generated normal subgroup of finite index which contains no element* $\neq \mathrm{id}$ *of finite order*.

For a proof see [Matsuzaki and Taniguchi 1998] or [Ratcliffe 1994].

To obtain the corresponding result for a finitely generated kleinian group, choose a set of N generators, and then pass to the matrix group generated by the $2N$ pairs of matrices $\{\pm A_i\}$.

Let G be a group generated by elements $g_1, g_2, \ldots$; we write $G = \langle g_1, g_2, \ldots \rangle$. A *word* in the chosen generators is a finite sequence (of length ≥ 0) whose elements are of the form g_i or g_i^{-1}; any such word gives rise, by multiplication, to an element of G. A word (of length > 0) giving rise to the identity of G is called a *relator* in G; it is called a *trivial relator* if it of the form $g_i g_i^{-1}$ or $g_i^{-1} g_i$ for one of the chosen generators.

Suppose $R_1, R_2, \ldots$ are relators in G. A word W is *derivable* from the relators $\{R_i\}$ if repeated application of the following operations changes W to the empty word in finitely many steps: Insertion or deletion of one of the words $R_1, R_1^{-1}, \ldots$, or of one of the trivial relators, between any two consecutive letters of W, or before or after the word W. If every relator is so derivable from the relators on the list $R_1, R_2, \ldots$ (plus the empty word), we say that the generators $g_1, g_2, \ldots$ and the relators $R_1, R_2, \ldots$ constitute a *presentation* of G, and we write

$$G = \langle g_1, g_2, \ldots \mid R_1, R_2, \ldots \rangle.$$

A *free group* is one that has a presentation $\langle g_1, g_2, \ldots \mid \rangle$; that is, if (for appropriately chosen generators) there are no nontrivial relators. A group is *finitely presented* if it has a presentation where both the generators g_i and the relators R_i are finite in number.

Suppose G is generated by $g_1, \ldots, g_N$ and $F_N = \langle f_1, \ldots, f_N \mid \rangle$ is a free group in the same number of generators. The map $\phi : f_i \to g_i$ extends to a homomorphism $\phi : F_N \to G$, and the elements in the kernel of ϕ are exactly the relators of G; any generating set for this kernel is a set of relators for a presentation of G.

The fundamental group of a closed surface of genus g has a presentation

$$\left\langle a_1, b_1, a_2, b_2, \ldots, a_g, b_g \mid \prod_{i=1}^{g}[a_i, b_i]\right\rangle,$$

where the generators come from appropriately chosen loops, as in Figure 2.1, p. 73.

For example, a closed surface of genus g has the single relation $\Pi_{i=1}^{g}[a_i, b_i] = 1$. Here $[a, b]$ denotes the commutator $aba^{-1}b^{-1}$. On the other hand, the fundamental group of a closed surface with punctures is a free group.

It is a basic property of hyperbolic 3-manifolds (and 3-manifolds more generally) that finitely generated fundamental groups are automatically finitely presented. The proof, due to Scott and Shalen, is a formal consequence of the existence of a compact core in the quotient manifold (see Section 3.9).

For orbifolds, Selberg's lemma can be applied.

Theorem 2.5.3 (Scott and Shalen). *Finitely generated kleinian groups are finitely presented.*

The finite presentation property is automatically true for compact manifolds, as we will see in Section 3.5.

This is in sharp contrast to the case of 4-manifolds, where any countable group, finitely presented or not, can be a fundamental group. There even exist finitely presented groups which have finitely generated subgroups which are not finitely presented [Scott 1973b].

2.6 Introduction to Riemann surfaces and their uniformization

A Riemann surface is a 1-dimensional complex analytic manifold: It is defined by coordinate coverings $\{U_\alpha, \phi_\alpha\}$ (of a connected Hausdorff space), where $\phi_\alpha : U_\alpha \to \mathbb{C}$ is such that the transition mappings $\phi_\beta {\phi_\alpha}^{-1}$ associated with overlapping coordinate neighborhoods are analytic homeomorphisms (conformal mappings).* Riemann surfaces are orientable and have countable bases. A homeomorphism $f : R \to S$ between Riemann surfaces is a *conformal mapping* if $\phi_\beta f {\phi_\alpha}^{-1}$ is conformal where defined. Usually one does not distinguish between conformally equivalent surfaces. The classic reference is [Ahlfors and Sario 1960], while [Farkas and Kra 1980] is excellent for closed surfaces, and the most elementary is [Springer 1957].

To put it informally, but more to the point, a Riemann surface is an oriented 2-dimensional surface with a rule (typically coming from a riemannian metric) for measuring angles; the angles about each point must sum to 2π. Two surfaces are

* If the transition mappings are instead required to be the restriction of Möbius transformations, the additional structure is called a complex projective structure. These structures will be explored in Exercise 6-14.

conformally equivalent if there is an orientation preserving homeomorphism that preserves angles as measured by the corresponding rules.

The most familiar cases are regions $\Omega \subset \mathbb{C}$ where the euclidean angles are taken. Make a new rule for measuring angles by defining the angle between two rays at $z \in \Omega$ to be the angle resulting after applying the affine transformation $T : (x, y) \mapsto (x', y')$ with $x' = x$, $y' = 2y$. This determines a new Riemann surface structure on the same underlying point set. However Ω, with its new structure, is conformally equivalent to $T(\Omega)$ with the natural structure from $\mathbb{C}$.

Another common situation is a smoothly embedded surface in $\mathbb{R}^3$ with the "rule" that is induced by the ambient euclidean metric. So is a polyhedral surface with the rule given by the euclidean metric in the polygons, except the neighborhood around each vertex must be flattened out so the angles add to 2π (the vertices can also be viewed as cone points with ,the cone angle being the sum of the vertex angles of the triangles sharing the vertex). In this connection, a mention of the following theorem [Rüedy 1971] is irresistible: Any abstract Riemann surface can be conformally embedded as a C^∞ surface, or even a polyhedral surface, in $\mathbb{R}^3$. Here the angles on the embedded surface are to be measured by restricting the ambient euclidean metric. If the surface is not compact, the ends of the embedded conformal equivalent go off to ∞. A conformal embedding can be found in an arbitrarily small neighborhood of a smoothly embedded model surface by deforming it in the normal direction.

However for most applications one works with Riemann surfaces that are not naturally embedded in any ambient space. Such an example is given below in terms of algebraic curves.

A Riemann surface may be of any genus $g \geq 0$ (number of "handles"), and with any number of "ends" (or "ideal boundary components"), countable or uncountable — like the Riemann surface which is the complement of the Cantor set. A *puncture* is an isolated ideal boundary component which has a neighborhood conformally equivalent to the once punctured unit disk. To put it another way, a *puncture* is obtained by removing a point from a Riemann surface. One can also speak of Riemann surfaces with borders — like the closed unit disk — but we will not be using them here.

A regular (unbranched) covering surface $\widetilde{R}$ of a Riemann surface is also a Riemann surface. The local complex structure can just be lifted. Deck transformations automatically become conformal automorphisms of $\widetilde{R}$.

A branched cover $\widetilde{R}$ is also a Riemann surface. It has a discrete set of special points called branch points. If $\xi^* \in \widetilde{R}$ is a branch point of order $r \geq 2$ and $\pi(\xi^*) = \xi$ is its projection in R, then given a small V neighborhood of ξ there is a neighborhood U of ξ^* such that $\pi(U) = V$ and each point $\neq \xi$ of V is covered exactly r times in U. If the branch values are removed from R and preimages from $\widetilde{R}$, one is left with a regular covering, that can be described by a subgroup of the fundamental group of the base surface.

For a survey of recent work on Riemann surfaces with singular conformal metrics see [Bonk 2002].

Conversely, if G is a discontinuous group acting on R then R/G is also a Riemann surface with R a possibly branched cover, depending on whether G has fixed points in R. A typical example is $\mathbb{H}^2/G$.

What is uniformization?

An abstract Riemann surface is described only in terms of local coordinates. Wouldn't it be nice if there were a global coordinate system $w = \phi(t)$, in terms of a complex parameter t, that served *uniformly* at all points? By way of analogy, the unit circle $\{w : |w| = 1\}$ is uniformized by the real line via the projection map $w = e^{it}$, $-\infty < t < \infty$.

For example, suppose P is an irreducible polynomial of two complex variables. Then $R = \{(x, y) \in \mathbb{C}^2 : P(x, y) = 0\}$ is a Riemann surface. To suggest why, suppose m is the degree of P in y. For most $x \in \mathbb{S}^2$ there will be m distinct values $y_k(x)$ that satisfy $P(x, y) = 0$; the m points $(x, y_k(x)) \in R$ lie over x. A small neighborhood N about such an x determines m disjoint neighborhoods $N_k \subset R$ and the map $x : (x, y) \mapsto x$ is a homeomorphism of each back down to N. The complex structure can be extended over the other points as well. As a result, R is a closed Riemann surface. Conversely, it is a famous classical theorem that every closed Riemann surface can be generated in this fashion.

A noteworthy class of examples are the *Fermat curves* $x^n + y^n = 1$, $n \geq 2$, which represent closed Riemann surfaces of genus $\frac{1}{2}(n-1)(n-2)$. The world now knows that when $n \geq 3$, there are no solution pairs of nonzero rational numbers. For algebraic curves more generally, Mordell's Conjecture is known to be true too: For curves $P(x, y) = 0$ of genus at least two, there are at most a finite number of solution pairs (x, y) where both x and y are rational numbers.

In short, for closed Riemann surfaces in particular, it would be nice if we could find a single complex parameter t such that $x = x(t)$, $y = y(t)$ for all points $(x, y) \in R$, that is, for all solution pairs of an associated $P = 0$.

Uniformization Theorem A. *A simply connected Riemann surface can be conformally mapped onto exactly one of*: *the Riemann sphere* $\mathbb{S}^2$, *the complex plane* $\mathbb{C}$, *the unit disk* $\mathbb{D}$.

Ahlfors [1973, p. 136] referred to the Uniformization Theorem as "the single most important theorem in the whole theory of analytic functions of one variable" (a proof can be found in that same reference). For regions properly embedded in $\mathbb{C}$ it reduces to the Riemann Mapping Theorem. The famous application is to the universal covering surface $\widetilde{R}$ of a Riemann surface R, which as defined is not embedded anywhere.

Uniformization Theorem B. *The universal cover* $\widetilde{R}$ *is conformally equivalent to*

(i) $\mathbb{S}^2$ *if and only if R is itself conformally* $\mathbb{S}^2$,
(ii) $\mathbb{C}$ *if and only if R is conformally equivalent to* $\mathbb{C}$, *to* $\mathbb{C} \setminus \{0\}$, *or to a torus*,
(iii) $\mathbb{D}$ *in all other cases*.

The group Γ of deck (or cover) transformations is isomorphic to the fundamental group $\pi_1(R)$. Deck transformations are conformal automorphisms of the universal covering, that is, Möbius transformations when one of the standard models are used. A deck transformation cannot have a fixed point in the cover, hence cannot be elliptic. Furthermore, Γ is *properly discontinuous* in $\widetilde{R}$: $N \cap \gamma(N) = \varnothing$ for any $\gamma \in \Gamma$ distinct from id, and any small neighborhood N of any point in $\widetilde{R}$. So the deck transformations form a discrete group.

Parabolic deck transformations are associated with *punctures* on R: recall that a puncture is an isolated "ideal boundary component" with the property that it has a "neighborhood" in R conformally equivalent to the once punctured unit disk. The lift of a small loop surrounding a puncture determines a parabolic transformation, and conversely, every parabolic transformation is associated with a puncture in this manner. If R has no punctures then Γ contains only loxodromic transformations (plus the identity).

Once we know that the abstract $\widetilde{R}$ is conformally equivalent to, say, the concrete $\mathbb{D}$, we can replace it by (identify it with) $\mathbb{D}$. Likewise, R is conformally equivalent to — and we can replace it by — the quotient surface $\mathbb{D}/\Gamma$ with Γ the group of deck transformations. The complex structure descends automatically from $\mathbb{D}$ to $\mathbb{D}/\Gamma$. The coordinate coverings are just $\{U_\alpha, z\}$, where the $U_\alpha \subset \mathbb{D}$ are small enough to project injectively, via the identity map, into R. In fact the group Γ is uniquely determined up to conjugation: if $\mathbb{D}/\Gamma$ is conformally equivalent to $\mathbb{D}/\Gamma_1$, then $\Gamma_1 = T\Gamma T^{-1}$ for some $\mathbb{D}$-preserving Möbius transformation T.

The third case of Theorem B is operative in particular whenever R is a closed Riemann surface of genus exceeding one. Such Riemann surfaces are algebraic curves, nonelliptic ones. Yet uniformization is not an algebraic process. One of the mysteries concerns the precise relation between an explicit polynomial that generates the surface and the uniformizing function.

The group of conformal automorphisms of a Riemann surface is discrete if and only if its fundamental group is nonabelian. For the record, the group of conformal automorphisms of a closed surface of genus $g \geq 2$ has at most $84(g-1)$ elements; see Exercise 3-1. The lowest genus for which this number can be attained is $g = 3$ and the surface that attains it is the *Klein surface*. In $\mathbb{D}$ it can be represented by fitting together 24 isometric regular hyperbolic heptagons with interior angles $2\pi/3$ (and area $\pi/3$); see Figure 2-5, on p. 87. After making the appropriate pairwise identification of the consequent free edges, the configuration "rolls up" to form a closed surface composed of 24 regular heptagons arranged in triples around 56 vertices. A model was sculpted by Helaman Ferguson and is on the terrace at the Mathematical Sciences Research Institute in Berkeley. The Klein surface as an algebraic curve is $x^3y + y^3 + x = 0$.

Even though it does not include the cases of simply and doubly connected plane regions, often one proclaims:

Hyperbolization Theorem for Riemann Surfaces. *Every Riemann surface with a*

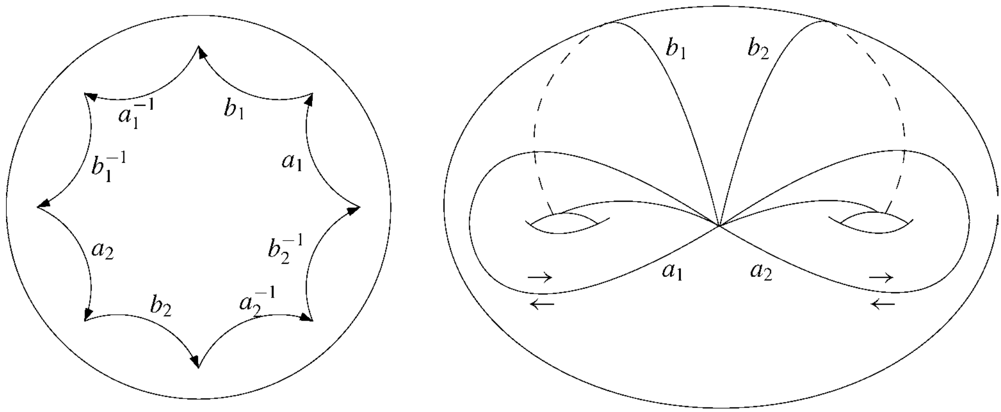


Fig. 2.1. Rolling up a regular octagon: The four transformations mapping edges $a_i \to a_i^{-1}$, $b_i \to b_i^{-1}$ and sending the octagon into its exterior generate the fuchsian covering group.

nonabelian fundamental group carries a hyperbolic metric compatible with its complex structure.

The Uniformization/Hyperbolization Theorem expresses the fact that when the universal cover $\tilde{R}$ can be taken as $\mathbb{D}$, we can bring to bear the dual role of $\mathbb{D} = \mathbb{H}^2$ having both a complex structure and a hyperbolic structure. The group Γ of deck transformations consists of conformal automorphisms of $\mathbb{D}$; therefore the complex structure of $\mathbb{D}$ induces a complex structure on $R = \mathbb{D}/\Gamma$. The group Γ is also a group of isometries of $\mathbb{H}^2$. Therefore the hyperbolic structure on $\mathbb{D} = \mathbb{H}^2$ induces a hyperbolic structure on $R = \mathbb{H}^2/\Gamma$. That is, if z denotes the coordinate in $\mathbb{D}$ and $w = \pi(z)$ the corresponding coordinate in R, define the hyperbolic metric on R by the equation $\lambda(w)\,|dw| = \rho(z)\,|dz|$, where $\rho\,|dz|$ is the hyperbolic metric in $\mathbb{D}$. Usually it is not possible to compute $\lambda(w) = \rho(z)/|\pi'(z)|$ with $z = \pi^{-1}(w)$ explicitly. Notable exceptions are the once punctured disk $\{0 < |w| < 1\}$ and annulus $\{1 < |w| < R\}$— see Exercise 2-2.

The surface R has finite hyperbolic area if and only if it is a closed Riemann surface of genus $g \geq 0$ with $n \geq 0$ points removed (punctures) satisfying $2g + n \geq 3$. Its area is $2\pi(2g + n - 2)$. Examples include the n-punctured spheres when $n \geq 3$.

On the one hand we can study analytic and meromorphic functions on R, in terms of its complex structure. On the other hand we can do geometry on R, talking about geodesics, triangles, etc. It is often easier to study the analysis and geometry in the universal cover, taking account of the covering group Γ. As explained in Section 3.5, there is a concrete model of R within $\mathbb{D}$ as a convex hyperbolic polygon called a Dirichlet region. Its sides are organized in pairs; when the polygon is "rolled up" by identifying the side pairs by Γ, a surface results and it is conformally equivalent to R.

There is a generalization of uniformization theory to Riemann surfaces with a discrete set of points designated as *cone points*. Assign to each of these points a rational cone angle of the form $2\pi/r$, where r is a positive integer ≥ 2. The choice $r = \infty$ means that the point should become a puncture. One requires now a *branched simply connected cover* with the following property. If ξ is a cone point with cone angle $\frac{2\pi}{r}$ then at each point ξ^* over ξ, the stabilizer of ξ^* in the cover group G is generated by an elliptic transformation of order r. A branched, simply connected covering corresponding to the assigned data exists as $\mathbb{S}^2$, $\mathbb{C}$, or most commonly $\mathbb{H}^2$ according to the possibilities described in Exercise 3-1.

We emphasize there are two aspects to the consideration of branch points. Consider the cyclic group $H = \langle z \mapsto e^{2\pi i/6} z \rangle$. A fundamental region for H in $\mathbb{D}$ is the sector $\{z : 0 \leq \arg z < 2\pi/6\}$. There are two ways to consider the quotient $R = \mathbb{D}/\langle z \mapsto e^{2\pi i/6} \rangle$. One way is require R to be a Riemann surface; necessarily then $R = \mathbb{D}$ and $\mathbb{D}$ is a branched covering of itself with projection map $w = z^6$. From a different prospective, R can be viewed as portion of the cone with cone angle $2\pi/6$ obtained when the fundamental sector $\{z \in \mathbb{D} : 0 \leq \arg z \leq 2\pi/6\}$ is rolled up to identify the edges. To make the cone into a Riemann surface at the cone point, it must be flattened out there. This is what is done by interpreting the map $z \mapsto z^6$ as a homeomorphism of the sector of central angle $2\pi/6$ with the edge identifications onto the full disk $\mathbb{D}$.

To a complex analyst, $\mathbb{D}/H$ is made into a Riemann surface by defining the complex structure in $\mathbb{D}/H$ in terms of the map $w = z^6 : \mathbb{D}/H \to \mathbb{D}$. On the other hand, a geometer sees $\mathbb{D}/H$ as the cone obtained by rolling up the sector, without bothering to define a complex structure at the cone point. The point $\{z = 0\}$ is called a *cone point*. The situation is analogous to that encountered by 3-dimensional orbifolds. An orbifold is actually a manifold, but that involves "flattening out" the cone points which is not such a natural operation.

The Uniformization Theorem can be divided into a topological part and an analytic part. The topological part says in particular that every orientable surface with non-abelian fundamental group has a hyperbolic structure — that is, it is homeomorphic to $\mathbb{H}^2/G$ for some fuchsian group G. This can be proven directly by modeling each surface type by a fuchsian group. The analytic part says that for a Riemann surface, the hyperbolic metric can be taken to be compatible with the conformal metric. It is the topological part that has an analogue for 3-manifolds, as is realized in the Hyperbolization Theorem, page 324. This too is proved by finding geometric models.

2.7 Fuchsian and Schottky groups

Fuchsian groups

Suppose G is a nonelementary, discrete group preserving the upper UHP and lower half-plane LHP. Each element $A \in G$ is symmetric in $\mathbb{R}$; it satisfies $\overline{A(z)} = A(\bar{z})$. Each elliptic transformation has one fixed point in UHP and the other at the symmetric point in LHP. The limit set is contained in $\mathbb{R}$. Classically, G is said to be of the *first kind* if $\Lambda(G) = \mathbb{R}$, otherwise it is said to be of the *second kind*.

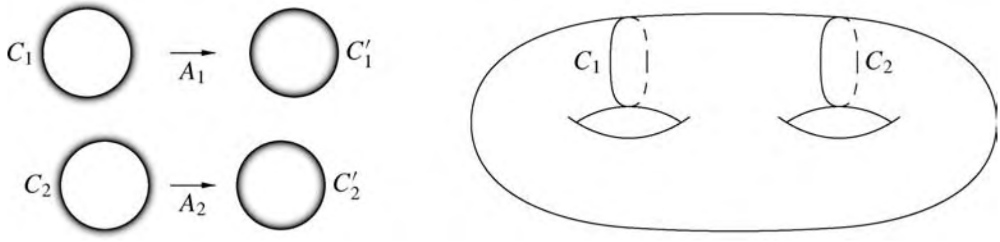


Fig. 2.2. A Schottky group's generators (left) and the group's quotient (right).

Suppose G is of the first kind. Then $\Omega(G) = \text{UHP} \cup \text{LHP}$ and if G is finitely generated, $R_{\text{top}} = \text{UHP}/G$ is a closed surface with at most a finite number of punctures and a finite number of branch values (cone points); see [Marden 1967; Casson and Bleiler 1988]. The quotients of the upper and lower half-plane are symmetric surfaces under reflection $z \mapsto \bar{z}$. The 3-manifold $\mathcal{M}(G) = \mathbb{H}^3 \cup (\text{UHP} \cup \text{LHP})/G$ is homeomorphic to $R_{\text{top}} \times [0, \pi]$; that is $\mathcal{M}(G)$ is an "I-bundle" with top surface R_{top}. This can be seen explicitly as follows. Let H_θ be the euclidean half-plane bordering $\mathbb{R}$, inclined at angle $0 < \theta < \pi$ to $\mathbb{C}$. G maps each half-plane H_θ onto itself. Their quotients H_θ/G are the cross sections in the I-bundle. In addition the orientation reversing involution $z \mapsto \bar{z}$ extends to all $\mathbb{H}^3$ and projects to an orientation reversing involution of $\mathcal{M}(G)$, interchanging its top and bottom boundary components, pointwise fixing the middle surface $R_{\pi/2}$.

Suppose instead G is of the second kind and nonelementary. Its ordinary set $\Omega(G)$ contains the countable number of open intervals $\Lambda_+ = (\mathbb{R} \cup \{\infty\}) \setminus \Lambda(G)$ and is connected and infinitely connected. The quotient $R_{\text{top}} = \text{UHP} \cup \Lambda_+/G$, if G is finitely generated, is a *compact bordered* Riemann surface containing at most a finite number of punctures and branch values. Its boundary consists of the finite number of simple closed curves Λ_+/G which are pointwise fixed by the involution of $\Omega(G)$, $z \mapsto \bar{z}$. Equally we have an involution of the surface $\Omega(G)/G$, which is called the *double* of R_{top}. The 3-manifold $\mathcal{M}(G)$ has a connected boundary and the product structure $R_{\text{top}} \times [0, 1]$, where the "top" and "bottom" pieces are joined across ∂R_{top}.

Schottky groups

This is the simplest class of function groups. Take $g \geq 1$ pairs of mutually disjoint circles in $\mathbb{C}$, $\{C_1, C_1', \dots, C_g, C_g'\}$, with mutually disjoint interiors. For each index, choose any Möbius transformation A_i that maps C_i to its partner C_i' and sends the interior of C_i to the exterior of its partner. The group generated by $\{A_i\}$ is called a *Schottky group of genus g*. It is the archetypical free group on g generators. The G-orbit of the circles nest down on the limit set $\Lambda(G)$ which is totally disconnected (every component of $\Lambda(G)$ is a point) as shown in Figure 2.7 (page 76). If G is not cyclic (which is an exceptionally simple special case), the limit set is a perfect set. In Mandelbrot's terminology it is "fractal dust", since it is known to have a positive Hausdorff dimension (Exercise 3-20). The ordinary set $\Omega(G)$ is connected and infinitely connected. The quotient surface $R = \Omega(G)/G$ is a closed surface of

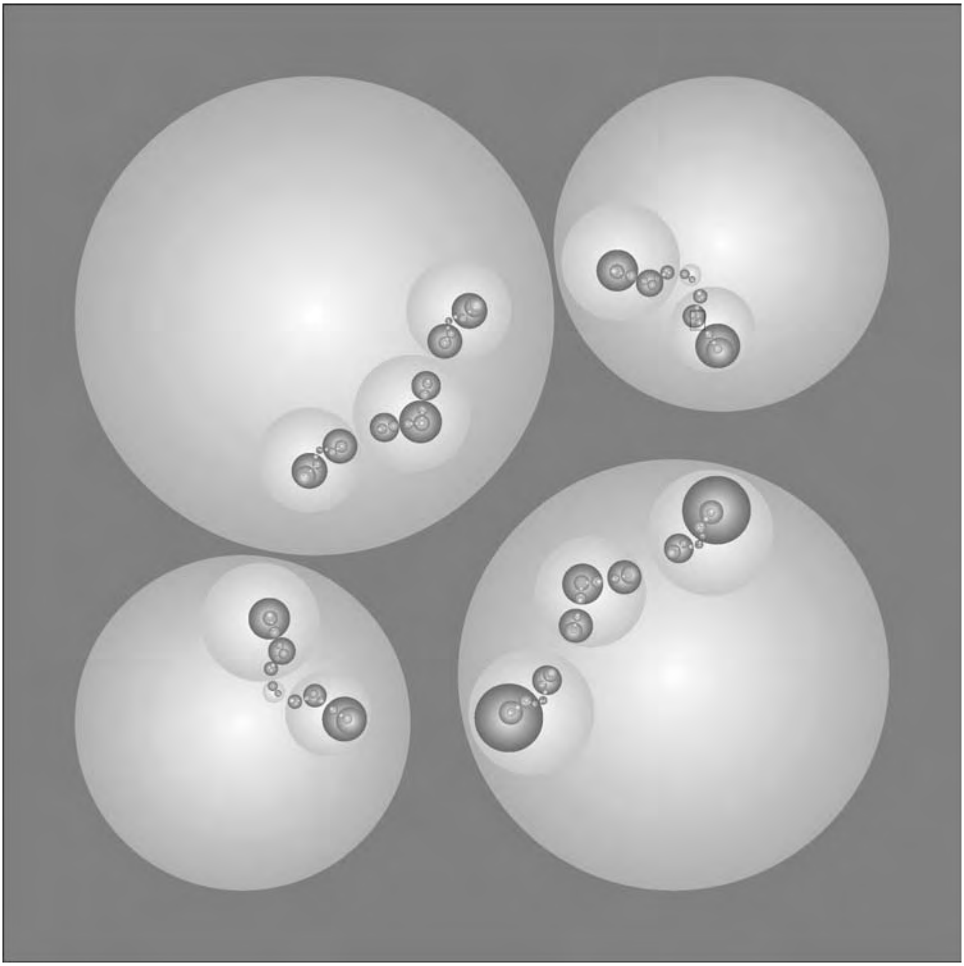


Fig. 2.3. A 2-generator Schottky group showing the orbit of the Schottky circles nesting to the limit set.

genus g. From the point of view of R, $\Omega(G)$ is a *planar covering surface*. There is a wonderful, thorough discussion of the two-generator case in [Mumford et al. 2002].

In the opposite direction, we have:

Maskit Planarity Theorem [Maskit 1988]. *Suppose R is a closed Riemann surface with at most a finite number of punctures and the covering surface $\widehat{R}$ determined by a normal subgroup N of $\pi_1(R; O)$ is planar. Then there is a finite set of mutually disjoint simple loops $\{\alpha_i\}$ in R and a corresponding set of integers $\{r_i \geq 1\}$ with the following property: N is the smallest normal subgroup* of $\pi_i(R; O)$ determined by $\{\alpha_i{}^{r_i}\}$, or equivalently, $\widehat{R}$ is the highest normal covering surface of R with the property that all lifts of the curves $\{\alpha_i{}^{r_i}\}$ are simple loops.*

* That is, N is generated by $\{\gamma\alpha_i{}^{r_i}\gamma^{-1}\}$ for all $\gamma \in \pi_1(R; O)$ with each α_i joined to O by an auxiliary path.

A *planar* Riemann surface $\widehat{R}$ is one which is conformally equivalent to a region in $\mathbb{C}$. Because the covering $\widehat{R} \to R$ corresponds to a normal subgroup $N \subset \pi_1(R)$, the deck transformations are conformal automorphisms of $\widehat{R} \subset \mathbb{C}$. If the result of cutting R along the simple loops $\{\alpha_i\}$ is itself a planar surface, then the deck transformations of $\widehat{R}$ are known to consist of the restrictions of Möbius transformations — see [Ahlfors and Sario 1960, IV.4B, IV.19F]. Otherwise, as shown in [Maskit 1968], there is a conformal map of $\widehat{R}$ onto another representation $\widehat{R}'$ of the covering for which the deck transformations become restrictions of Möbius transformations (Exercise 2-16).

Returning to Schottky groups, the quotient manifold $\mathcal{M}(G)$ is a *handlebody of genus* g. The common exterior of the circles serves as a fundamental region for $\Omega(G)$. The common exterior of the hyperbolic planes rising from the circles serves as a fundamental region in $\mathbb{H}^3$. For $g = 1$ it is a *solid torus* — a bagel! More generally $\mathcal{M}(G)$ is homeomorphic to the result of gluing g bagels together.

A handlebody M of genus $g \geq 1$ is characterized by the following property. Its boundary ∂M is a closed surface of genus g. There exist g mutually disjoint simple curves on ∂M called *compressing curves*, each of which bounds a disk within M — in our construction above these disks can be taken to be the planes rising from the circles — such that when M is cut along these disks what results is connected and homeomorphic to a ball. If the handlebody M is embedded in $\mathbb{R}^3$, its exterior in $\mathbb{S}^3$ is either a handlebody, with its own, distinct, collection of compressing curves on ∂M, or it is knotted.

Suppose X_1 and X_2 are two handlebodies of the same genus and $\Phi : \partial X_1 \to \partial X_2$ is a homeomorphism. Attach X_1 to X_2 by identifying each point $x \in \partial X_1$ to $\Phi(x) \in \partial X_2$. The result is a closed orientable 3-manifold M. Conversely it has long been known that in *every* closed, orientable 3-manifold M^3, one can find embedded surfaces S with the property that $M^3 \setminus S$ is the union of two handlebodies; see [Hempel 1976] or [Jaco 1980]. Such a decomposition is called *Heegaard splitting*. For example a Heegaard splitting of a closed manifold can be obtained by taking a tubular neighborhood about the union of 1-simplices of a triangulation. There are two sets of simple loops $\mathcal{S}_1$, $\mathcal{S}_2$ on S such that each loop in $\mathcal{S}_1$ bounds a disk in one component of $M^3 \setminus S$ and each loop in $\mathcal{S}_2$ bounds a disk in the other. There is an orientation reversing homeomorphism between the two components that interchanges the two sets. The seeming simplicity of the splitting is very deceptive; all efforts to decipher the topology of the resulting manifold from the homeomorphism Φ and its interplay with compressing curves have failed.

It is a conjecture of Agol that every 2-generator, closed kleinian manifold is homeomorphic to the result of so gluing two genus-2 handlebodies.

The Schottky construction works equally well if we replace the circle pairs by pairs of Jordan curves that are known to be associated with Möbius transformations sending the interior of one to the exterior of its partner. To reflect this distinction the Schottky groups generated by circles are known to aficionados as *classical Schottky groups*, whereas groups with the less restrictive requirement are known merely as

Schottky groups. The more general situations arise naturally in planar uniformizations of surfaces. It is known [Marden 1974c] that not every Schottky group in the general sense can be generated by circles, no matter how the generators are chosen — for examples see [Gilman and Waterman 2003]. In any case Schottky groups form that class of kleinian groups for which $\mathcal{M}(G)$ is a handlebody. The handlebodies of genus g obtained from classical groups are characterized by the property of containing g mutual disjoint hyperbolic planes that are bounded by simple loops which are not retractable to points in the boundary.

Conversely, a discrete, finitely generated, purely loxodromic group with $\Omega(G) \neq \varnothing$ that is a free group is automatically a Schottky group [Maskit 1988, X.H.6]. A special case is a finitely generated fuchsian group of the second kind, without elliptics or parabolics. In fact this is a classical Schottky group; it is an illuminating exercise to verify this fact directly. See Exercise 2-18.

2.8 Riemannian metrics and quasiconformal mappings

In terms of local coordinates, a smooth, nonsingular riemannian metric $ds^2 = E\,dx^2 + 2F\,dx\,dy + G\,dy^2$ on surface element (or a region $\Omega \subset \mathbb{C}$) can be written in complex form in terms of $z = x+iy$, $\bar{z} = x-iy$ as $ds^2 = \lambda(z)|\,dz + \mu(z)\,d\bar{z}|^2$, where $\lambda(z) > 0$ and $0 \le |\mu(z)| \le k < 1$. In the special case $\mu = 0$, it is a *conformal metric* $|dw| = \lambda(z)|dz|$; this means that a tiny circle $|z| = \varepsilon$ will become a tiny circle $|w| = \lambda(0)\varepsilon$ and the angle measure is the same as in the z-coordinate.

Given the riemannian metric, we can introduce new local coordinates — we can change the rule for measuring angles — on the surface in terms of which the metric becomes conformal. This is a classical procedure called introducing *isothermal coordinates*. Put another way, changing to isothermal coordinates makes the surface into a Riemann surface. Or, if we start with a Riemann surface and a riemannian metric, the metric determines a new Riemann surface structure on the underlying pointset.

To find the new structure, we must solve in each coordinate patch the *Beltrami equation*

$$\frac{\partial F}{\partial \bar{z}} = \mu(z)\frac{\partial F}{\partial z}. \tag{2.4}$$

If $\mu = 0$, (2.4) reduces to the Cauchy–Riemann equations for analyticity. A solution F will satisfy the infinitesimal equation $|dF| = |F_{\bar{z}}|\,|dz + \mu(z)\,d\bar{z}|$. It is the solution $w = F(z)$ that is used to introduce a new complex structure — a new rule for measuring angles — on the same underlying pointset.

For example, consider the metric $ds^2 = |dz + k d\bar{z}|^2$ in $\mathbb{C}$. The map $w = F(z) = z + k\bar{z}$, $z \in \mathbb{C}$, $0 < k < 1$, solves the Beltrami equation with $\mu = k$. This is an orientation preserving (since $k < 1$), nonsingular (since $k \neq 1$), homeomorphism sending circles about $z = 0$ to ellipses with major and minor axes in the ratio $K = (1+k)/(1-k)$. Introduce a new angle measure at $z = 0$ by defining the angle between two rays to be the angle between the image of the rays. This is the angle measure determined by the riemannian metric $|dz + k\,d\bar{z}|^2$. In this case the new Riemann surface is also $\mathbb{C}$.

In the general theory, $\mu(z)$ needs only to be measurable on its domain, say $\Omega \subset \mathbb{C}$ with essential supremum $\|\mu\|_\infty = k < 1$. The Beltrami equation has a solution F which is a *K-quasiconformal mapping*.*† Near a point, say $z = 0$, where F is differentiable (which it is almost everywhere), F is approximated by an affine map $z \mapsto az + b\bar{z}$. A solution F is uniquely determined up to postcomposition with conformal mappings of its range. Indeed, if g is a conformal mapping of the range, then both F and $g \circ F$ satisfy the same Beltrami equation. The number $K = (1+k)/(1-k)$, where $k = \|\mu\|_\infty$, is called the *maximal dilatation* of F, and $\mu = F_{\bar{z}}/F_z$ is called its *complex dilatation*. The maximal dilatation measures the maximal distortion of the mapping in the sense that infinitesimal circles are sent to infinitesimal ellipses with ratio of major to minor axis uniformly bounded by K; $K = 1$ if and only if F is conformal.

The inverse of a K-quasiconformal mapping is also a K-quasiconformal mapping.

It is often better to define the complex dilatation μ on all $\mathbb{C}$ — which can be regarded as $\mathbb{S}^2$ as the values of μ at isolated points do not matter. For example, set $\mu = 0$ on the complement of Ω. When μ is defined on $\mathbb{S}^2$, except perhaps for a set of zero spherical area, and satisfies $\|\mu\|_\infty < 1$ then there is a unique solution of the Beltrami equation up to postcomposition by Möbius transformations. It is a homeomorphism $\mathbb{S}^2 \to \mathbb{S}^2$. Consequently the easiest way of normalizing the solution is to require that it fix three prescribed points.

Now suppose $\Omega \subset \mathbb{C}$ is preserved by a kleinian group G. We want to consider mappings F that project to map Ω/G onto itself or onto another Riemann surface. For this to happen, μ must be a *Beltrami differential* with respect to G, that is, μ must imply, for any $g \in G$, that both F and $F \circ g$ satisfy the same Beltrami equation. The condition that this be the case is

$$\mu(g(z))\frac{\overline{g'(z)}}{g'(z)} = \mu(z) \quad \text{for all } g \in G \text{ and (almost) all } z \in \Omega. \tag{2.5}$$

Often we will extend μ to $\mathbb{C}$ by setting it equal to zero in the complement of Ω, so μ will automatically become a Beltrami equation for G in $\mathbb{S}^2$. Any solution F will a quasiconformal map when restricted to Ω and a conformal mapping when restricted to the complement of $\overline{\Omega}$.

If we know that $\Lambda(G)$ has zero area, it suffices to require μ to be a Beltrami differential on $\Omega(G) = \mathbb{S}^2 \setminus \Lambda(G)$.* If μ satisfies (2.5) on $\mathbb{C}$, then both F and $F \circ g$ are solutions of (2.4) for any $g \in G$. Therefore, if F has been normalized, there exists a uniquely determined Möbius transformation $\varphi(g)$ with the property that $F \circ g(z) = \varphi(g) \circ F(z)$ for all $z \in \mathbb{S}^2$. Consequently F induces an isomorphism

* F is an orientation preserving homeomorphism with locally integrable distributional derivatives $F_z, F_{\bar{z}}$.

† The equivalent geometric definition that generalizes to arbitrary metric spaces is that a homeomorphism F of Ω is quasiconformal if there is some constant $H < \infty$ such that for every $z \in \Omega$,

$$\limsup_{r\to 0} \frac{\sup_{|w-z|=r} |f(w) - f(z)|}{\inf_{|w-z|=r} |f(w) - f(z)|} \le H.$$

* Actually we do not have to worry about $\Lambda(G)$ at all, only about the ordinary set, by Sullivan's Theorem (p. 158).

$\varphi : G \to \varphi(G) = H$ onto another kleinian group H. The group $H = FGF^{-1}$ is called a *quasiconformal deformation* of G. It is a *trivial deformation*, really not a deformation at all, if for some Möbius U, $\varphi(g) = UgU^{-1}$ for all $g \in G$, that is, if φ is a conjugation.

Suppose G is a fuchsian group acting in the upper half-plane UHP and μ is a Beltrami differential for G. There is a way of arranging things so that the quasiconformal deformation is fuchsian as well. This is done by extending μ by symmetry to the lower half-plane LHP: $\mu(z) = \overline{\mu(\bar{z})}$. It will remain a Beltrami differential for G. Normalize F so as to fix, for example, $(0, 1, \infty)$. Then F maps each of UHP and LHP onto itself. The quasiconformal deformation $H = FGF^{-1}$ is fuchsian.

Another example is a Schottky group G. Take a Beltrami differential μ in its ordinary set $\Omega(G)$. We don't need to bother with the limit set because it has zero area. A normalized solution F of the Beltrami equation will induce an isomorphism φ onto another Schottky group H. Even if G is a classical Schottky group it is unlikely that the F-images of the Schottky circles are round circles. But the pairing geometry of these F-images will remain the same.

Return to the case that Ω is simply connected. At the quotient level: F induces a quasiconformal mapping $f_* : R = \Omega/G \to S = F(\Omega)/H$. In each homotopy class $[f]$ of a quasiconformal map between the two surfaces, there will be uncountably many quasiconformal mappings. One of them may even be conformal. If so, the deformation, or deformation class $[f]$, is said to be *trivial*—there has been no real deformation at all. One of Teichmüller's basic contributions is a characterization of trivial classes.

Many people have tried to resolve the Ehrenpreis–Siegel conjecture: given any two closed Riemann surfaces R_1, R_2 and $\varepsilon > 0$, does there exist finite-sheeted, unbranched covers R_1^* of R_1 and R_2^* of R^* which are homeomorphic and are close to each other in the sense that there is a quasiconformal map $F : R_1^* \to R_2^*$ with complex dilatation $\|\mu\|_\infty < \varepsilon$?

Teichmüller spaces of Riemann surfaces

Suppose $R = \mathbb{H}^2/G$ is a closed Riemann surface of genus $g \geq 0$ with $n \geq 0$ punctures such that $3g + n - 3 > 0$. The *Teichmüller space* Teich(R) is defined as the quotient space

$$\mathrm{Teich}(R) = \{(S, f) \mid f : R \to S \text{ is quasiconformal}\}/\equiv,$$

with the equivalence

$$(S, f) \equiv (S', f') \text{ if and only if } f' \circ f^{-1} : S \to S' \text{ is homotopic to a conformal map.}$$

If R is a closed surface, we can use the term "orientation preserving homeomorphism" rather than "quasiconformal". The latter is needed only to insure that the punctures are not opened up to holes. We emphasize that Teich(R) is the space of "marked" Riemann surfaces: each equivalence class is associated with a particular homotopy class of maps $R \to S$, or an isomorphism $\pi_1(R) \to \pi_1(S)$, that relates each point to the basepoint (R, id).

A lift F to $\mathbb{H}^2$ of a quasiconformal mapping $f : R = \mathbb{H}^2/G \to S = \mathbb{H}^2/H$ has a homeomorphic extension to $\partial\mathbb{H}^2$. It induces an isomorphism $\theta : G \to H$. If $\ell \subset \mathbb{H}^2$ is a geodesic, the endpoints of $F(\ell)$, which in general is not a geodesic, are the F-images of the endpoints of ℓ. In other words, F determines a injection between geodesics on R and S.

Another definition is as the space of hyperbolic metrics (curvature -1) on a fixed say C^∞-surface R; instead of changing surfaces, change metrics: A hyperbolic metric is associated with each riemannian metric on R via isothermal coordinates and the uniformization theorem. An orientation preserving C^∞-diffeomorphism $h : R \to R$ sends the hyperbolic metric to another (set $z = h(w)$ in $|dz|/y$, for example). Let Diff_0 denote the group of those diffeomorphisms which are homotopic (and hence isotopic) to the identity.

$$\mathrm{Teich}(R) = \{g : g \text{ is a hyperbolic metric on } R\}/\mathrm{Diff}_0.$$

That is, two metrics are identified if they differ by a diffeomorphism homotopic to identity.

A third definition is as the deformation space of fuchsian groups:

$$\mathrm{Teich}(G) = \{\theta \mid \theta : G \to G' \text{ is a type preserving isomorphism to a fuchsian} G'\}/\equiv .$$

(*Type preserving* here means that parabolics correspond to parabolics.) Here θ corresponds to a homotopy class of quasiconformal maps $\mathbb{H}^2/G \to \mathbb{H}^2/G'$, and θ, θ' represent the same point if and only if they are conjugate: $\theta(g) = U \circ \theta'(g) \circ U^{-1}$ for some U and all $g \in G$.

Teichmüller's famous theorem tells us that among all the quasiconformal maps in the homotopy class of a quasiconformal map $g : S \to S_1$ there is a unique extremal mapping, called a *Teichmüller mapping*, that minimizes the maximal dilatation K among all quasiconformal mappings in the homotopy class. The *Teichmüller distance* between two points (S, f), $(S_1, f_1) \in \mathrm{Teich}(R)$ is defined as $\log K$, where $K = (1+k)/(1-k)$ is the minimal maximal dilatation of all quasiconformal mappings in the homotopy class $[f_1 \circ f^{-1} : S \to S_1]$: There is exactly one such mapping whose maximal dilatation achieves the value K. The theory [Strebel 1984] shows that each such extremal mapping F corresponds to a Beltrami equation of the following form: Corresponding to F is is a uniquely determined (up to positive constant multiple) holomorphic quadratic differential $\varphi_F(z)\,dz^2$ on S such that F and $\varphi = \varphi_F$ are related by the Beltrami equation $F_{\bar z} = k(\overline{\varphi}/|\varphi|)F_z$.

With this as metric, $\mathrm{Teich}(R)$ turns into a metric space, homeomorphic to $\mathbb{R}^{6g+2n-6}$. Each point lies on a uniquely determined geodesic ray from a given point S determined by the solution to a Beltrami equation $F_{\bar z} = t(\overline{\varphi}/|\varphi|)F_z$, for some $0 \le t < 1$ on S. See also Exercise 5-23.

If $\alpha : R \to R$ is a quasiconformal automorphism, then α — or rather its homotopy class — determines an automorphism, also denoted by α, of $\mathrm{Teich}(R)$ defined on equivalence classes by $\alpha : (S, f) \mapsto (S, f \circ \alpha)$. The totality of homotopy classes of such automorphisms $[\alpha]$ form the *mapping class group* (or *Teichmüller modular group*) $\mathfrak{M}(R)$*. In the case of the deformation space of tori, this corresponds to the classical modular group (Exercise 2-5). It is a celebrated theorem of Royden [1971] that $\mathfrak{M}(R)$ constitutes the full group of isometries of Teich(R).

A point (S, f) is fixed by $[\alpha] \in \mathfrak{M}(R)$ if and only if $f\alpha f^{-1} : S \to S$ is homotopic to a conformal map, that is if S has a conformal symmetry in the homotopy class of $f\alpha f^{-1}$. If α has a fixed point it has finite order — it can be thought of as an "elliptic" element. It is a famous theorem of Kerckhoff [1983], resolving a longstanding conjecture called the Nielsen Realization Problem: Corresponding to every *finite* subgroup $F \subset \mathfrak{M}(R)$, there exists $(S, f) \in$ Teich(R) such F corresponds to a finite group of conformal automorphisms of S; that is, F has a common fixed point in Teich(R).

Besides the elements of finite order, $\mathfrak{M}$ has elements analogous to the parabolics and loxodromics of kleinian groups. We will return to these matters in Exercise 5-6.

Rather than bothering with homotopy classes, one might wonder if $\mathfrak{M}(R)$ is isomorphic to an actual group of homeomorphisms of R, as it is in the case of a torus or in the case of a finite subgroup. The answer is negative for closed surfaces of genus exceeding five [Markovic 2005].

For closed surfaces of genus exceeding two, the mapping class group is generated by six elements, each of order two [Brendle and Farb 2004]. The mapping class group is also known to be finitely presented.

The mapping class group acts discontinuously on Teich(R). The quotient orbifold Teich$(R)/\mathfrak{M}(R)$ is called the *moduli space*. It comprises all Riemann surfaces which are quasiconformally equivalent to R; two such surfaces determine the same point if and only if they are conformally equivalent, no matter what homotopy class the conformal mapping is in. Unlike Teich(R), the moduli space is an algebraic object.

It is an interesting fact that in parallel to Selberg's Lemma for Möbius transformations, there is a torsion free subgroup $\mathfrak{M}_0 \subset \mathfrak{M}(G)$ of finite index, see [Ivanov 1992].

Teich(R) is also a complex analytic manifold of dimension $(3g + b - 3)$, where g is the genus of R and $b \geq 0$ is the number of punctures.* With respect to the analytic structure, the modular group constitutes the full group of biholomorphic automorphisms of Teich [Earle and Kra 1974]. Thus Teich$(G)/\mathfrak{M}(G)$ is an analytic orbifold and its finite sheeting covering Teich$(G)/\mathfrak{M}_0(R)$ is an analytic manifold. Unlike Teich(R), these quotients are open subsets of (compact) algebraic varieties.

We will present a "concrete" realization of Teich(R) in Section 5.6 in which its complex structure is more apparent.

For excellent introductions to the theory of quasiconformal mappings and Teich-

* As defined here, $\mathfrak{M}(R)$ consists of orientation preserving mappings; one can also consider the *extended mapping class group* which consists in addition of orientation reversing mappings.

* In contrast fuchsian representations of $\pi_1(R)$ depend on $6g + 2b - 6$ real numbers, including one relation costing 3 real numbers and normalization costing another 3.

müller spaces, we cite [Ahlfors 1966; Lehto 1987; Imayoshi and Taniguchi 1992].

2.9 Exercises and explorations

2-1. *Elementary and reducible groups.*

(i) Prove that a (not necessarily discrete) group G is elementary if and only if any two elements of infinite order have a common fixed point.

Hint: Clearly the definition given in Section 2.2 implies this one. If all elements are elliptic, we must appeal to the fact, to be proven in Corollary 4.1.5, that either the group is cyclic, or it has a common fixed point in $\mathbb{H}^3$. All the parabolic elements of G must share the same fixed point. If A and B share the fixed point ζ and at least one of them is not parabolic, then their commutator $[A, B] = ABA^{-1}B^{-1}$ is either parabolic or the identity; it is the identity if and only if A and B have the same set of fixed points. Thus for $B \in G$ with two fixed points, there cannot exist two other transformations, A, C such that A, B share one fixed point of B and C, B share the other. In particular all parabolic and loxodromic elements of G must have a common fixed point ζ. Then any elliptic element must fix ζ as well, unless all the loxodromic elements have the same pair of fixed points and the elliptic element interchanges the two fixed points.

Conclude as in Lemma 2.3.1(v) that every nonelementary group contains two loxodromic elements without a common fixed point.

(ii) A Möbius group H is called *reducible* if there is a fixed point common to all elements of H. A reducible group is, in particular, elementary.

Suppose that H is not reducible. Show that there exist two elements without a common fixed point. As a consequence show that a nonabelian group is reducible if and only if the trace of every commutator is $+2$ (Lemma 1.5.1).

Hint: Assume that H is not reducible. Choose an element $h \neq \mathrm{id}$. If h is parabolic, there is an element which does not fix the fixed point of h. Instead suppose h is loxodromic or elliptic with fixed points $\zeta_1, \zeta_2 \in \mathbb{S}^2$. If there is an element with distinct fixed points we are finished. Otherwise there is an element h_1 which fixes ζ_1 but not ζ_2, and h_2 which fixes ζ_2 but not ζ_1. If h_1 and h_2 have distinct fixed points we are done. Otherwise h_1 and h_2 have a common fixed point $\zeta_3 \neq \zeta_1, \zeta_2$. But $h_2 \circ h_1$ fixes neither ζ_1 nor ζ_2.

2-2. Show that $w = e^{2\pi i z}$ is a conformal mapping of the quotient space

$$\mathbb{H}^2/\langle z \mapsto z + 1\rangle$$

onto the punctured disk $0 < |w| < 1$. Find a corresponding mapping from

$$\mathbb{H}^2/\langle z \mapsto kz\rangle,$$

where $k>1$, to some annulus $1<|w|<R$. Then show that the hyperbolic metrics $\lambda(w)\,|dw|$ in the punctured disk and annulus are given by

$$\lambda(w)=\frac{1}{|w|\log\dfrac{1}{|w|}},\qquad \lambda(w)=\left(\frac{\pi}{\log R}\right)\frac{1}{|w|\sin\dfrac{\pi\log|w|}{\log R}}.$$

In the annulus, the geodesic is the circle $\{|w|=\sqrt{R}\}$, it is fixed by the involution, and its hyperbolic length is $\frac{2\pi^2}{\log R}$.

In contrast, verify the following formula for the mildly singular metric that results from pulling the hyperbolic metric down to $\mathbb{H}^2/\langle E\rangle$, where E is elliptic of order n:

$$\lambda(w)=\frac{2}{n|w|^{(n-1)/n}(1-|w|^2/n)}.$$

Hint: use the disk model and the map $w=z^n$.

2-3. Prove the result of C. L. Siegel repeated in [Lehner 1964, Theorem III.J] that a nonelementary group that preserves the upper half-plane and which is not discrete contains an elliptic element of arbitrarily high order.

Hint: Suppose $A=\left(\begin{smallmatrix}\lambda & 0\\ 0 & 1/\lambda\end{smallmatrix}\right)$ is an element of G and there is a sequence $\{B_n\}$ with $\lim B_n=\mathrm{id}$. Compute the trace of the commutators $C_n=AB_nA^{-1}B_n^{-1}$ and $D_n=AC_nA^{-1}C_n^{-1}$. Writing the normalized matrix $B_n=\left(\begin{smallmatrix}a_n & b_n\\ c_n & d_n\end{smallmatrix}\right)$, show first that $\lim b_nc_n=0$ so that $\lim a_nd_n=1$ and $a_nd_n>0$ for large indices. Conclude that for infinitely many indices either $\mathrm{tr}^2(C_n)<4$, so that C_n is elliptic, or $\mathrm{tr}^2(D_n)<4$.

2-4. Suppose $\{T_n\}$ is a sequence of loxodromic or elliptic transformations such that $(\mathrm{tr}\,T_n)^2$ has limit 4. Show that for a subsequence, there is a sequence of conjugates $\{U_kT_kU_k^{-1}\}$ such that $\lim U_kT_kU_k^{-1}$ is a parabolic transformation. One example is the sequence $\{z\mapsto e^{2\pi i/n}z\}$. (For an application, see the video *Not Knot* [Gunn and Maxwell 1991].)

2-5. *The modular group.* For this exercise it may be helpful to refer, for example, to [Ahlfors 1978]. The group M of normalized matrices with integer entries is called the *modular group* and is often denoted by $\mathrm{Mod}=\mathrm{PSL}(2,\mathbb{Z})$. This is an object of fundamental importance in number theory, in particular in the proof of Fermat's Last Theorem by Wiles. It is also involved in the theory of quadratic forms: if an integer N can be represented as $N=ax^2+2bxy+cy^2$, where a,b,c are given integers and x,y are integer variables, then replacing x,y by $mx+ny$, $px+qy$, where $\left(\begin{smallmatrix}m & n\\ p & q\end{smallmatrix}\right)\in\mathrm{Mod}$, gives a new representation of N.

Show that is generated by $z\mapsto z+1$ and $z\mapsto -1/z$. It can also be expressed as the free product $\mathbb{Z}_2*\mathbb{Z}_3$, that is, it is generated by elliptics of order two and three with parabolic commutator. Confirm that the following is a fundamental polygon F for the action of Mod in the upper half-plane: $\{z:-\frac{1}{2}<\mathrm{Re}\,z\le\frac{1}{2},\ |z|>1\}$ with the boundary segment $\{|z|=1,\ 0\le\mathrm{Re}\,z\le\frac{1}{2}\}$.

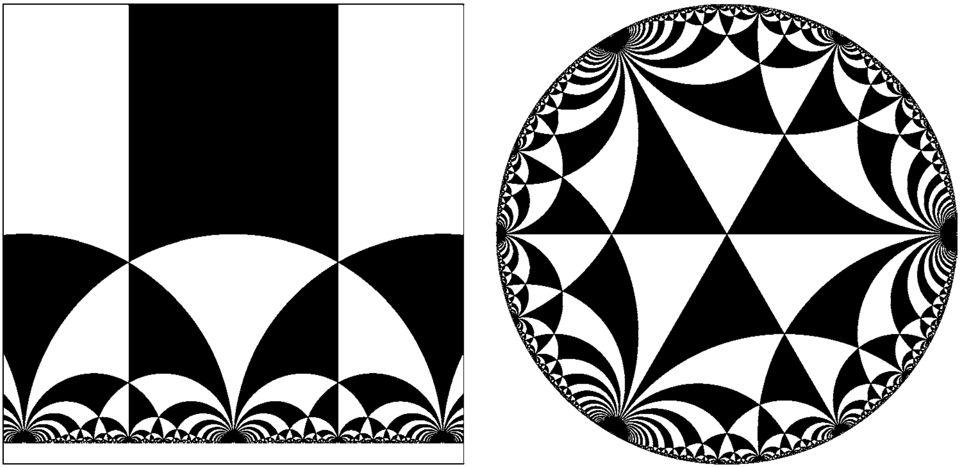


Fig. 2.4. Tessellations of the upper half-plane and disk by the orbits of the standard fundamental polygon under the modular group.

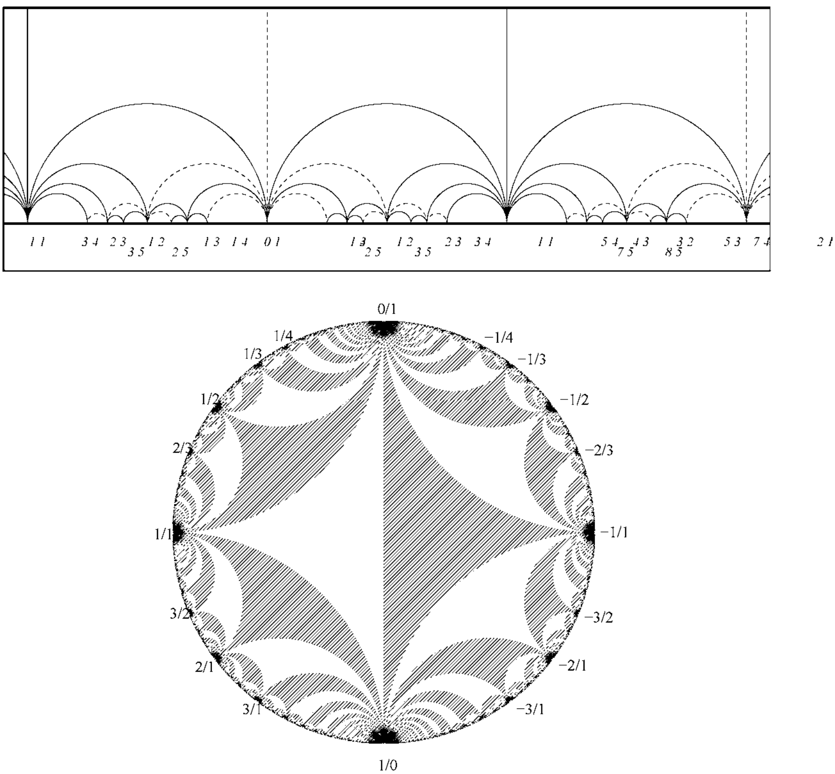


Fig. 2.5. Tessellation of the upper half-plane and disk by the orbit of an ideal quadrilateral, the union of two adjacent ideal triangles, under the 3-punctured sphere group. The ideal vertices are labeled by the Farey sequence (Exercise 2-9), which are their coordinates on $\mathbb{R}$. If the endpoints of the outer edge of an ideal triangle are $p/q < r/s$, then $ps - qr = \pm 1$ and the coordinate of the third vertex is $(p+r)/(q+s)$.

For later use also note that the intersection F_1 of the region in the upper half-plane $\{|z-n| \geq 1 : n \in \mathbb{Z}\}$ with the strip $\{0 < \operatorname{Re} z < 1\}$ also serves as a fundamental polygon for M (with due consideration for boundary arcs).

Denote by M_2 the subgroup of the modular group M, called the *level* 2 *congruence subgroup* of M, consisting of normalized matrices which satisfy

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bmod 2.$$

Show that M_2 is generated by $z \mapsto z+2$ and $z \mapsto z/(2z+1)$. Furthermore M_2 has index 6 in M; show this by showing that the ideal quadrilateral

$$F_2 = \left\{z : -1 < \operatorname{Re} z \leq 1,\ |z+\tfrac{1}{2}| > \tfrac{1}{2},\ |z-\tfrac{1}{2}| \geq \tfrac{1}{2}\right\}$$

is a fundamental polygon for M_2 and that it contains 6 copies of F. The quotient surface $\mathbb{H}^2/M_2$ is conformally equivalent to the triply punctured sphere.

Conversely suppose all three transformations $A, B, C = AB$ are parabolic with distinct fixed points. Prove that $\langle A, B\rangle$ is a discrete group preserving some round disk in $\mathbb{S}^2$. In fact, it is conjugate to M_2. In short, there is only one triply punctured sphere, up to Möbius equivalence.

More generally, a fuchsian group Γ is called a (hyperbolic) *triangle group* of signature (p, q, r), $2 \leq p, q, r \leq \infty$, if it is generated by elements A, B such that A, B, $C = BA$ are elliptic of orders p, q, r — or parabolic if the corresponding order is infinite. Providing $\frac{1}{p}+\frac{1}{q}+\frac{1}{r} < 1$, there exists such a fuchsian group (Exercise 3-1). The triangle group arises from a hyperbolic triangle with vertex angles $(\pi/p, \pi/q, \pi/r)$ by first taking the group $\langle a, b, c\rangle$ generated by the reflections in the sides and then passing to the index two, orientation preserving, subgroup generated by $A = ab$, $B = bc$, $C = ca$. The presentation is $\Gamma = \langle A, B, C : A^p = B^q = C^r = ABC = 1\rangle$.

The modular group has signature $(2, 3, \infty)$ and M_2 has signature (∞, ∞, ∞). Prove that up to Möbius equivalence there is only one group for each such signature. In fact, show using the trace identities that a triangle group must necessarily be fuchsian.

While we are dealing with the fundamental polygon for M_2 we will take the opportunity of pointing out the following phenomenon. If we move the fundamental polygon to the unit disk $\mathbb{D}$, it is bounded by a chain of four circular arcs orthogonal to $\partial\mathbb{D}$ and mutually tangent at their points of intersection. The group $\mathcal{M}_2$ is generated by pairing successive arcs, sending the exterior of one to the interior of its partner. The four points of tangency correspond to the three punctures on the quotient 3-punctured sphere. But we can equally pair the opposite arcs instead of the adjacent ones. Show that this results in a quotient which is a once punctured torus!

In particular we have shown that the same fundamental polygon can serve for two entirely different groups.

$\mathrm{SL}(2, \mathbb{Z})$ itself is generated by $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$, and also by $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $V = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$. It has the presentation $\langle A, B \mid A^2 = B^3, A^4 = \mathrm{id}\rangle$. See [Magnus 1974, p. 108].

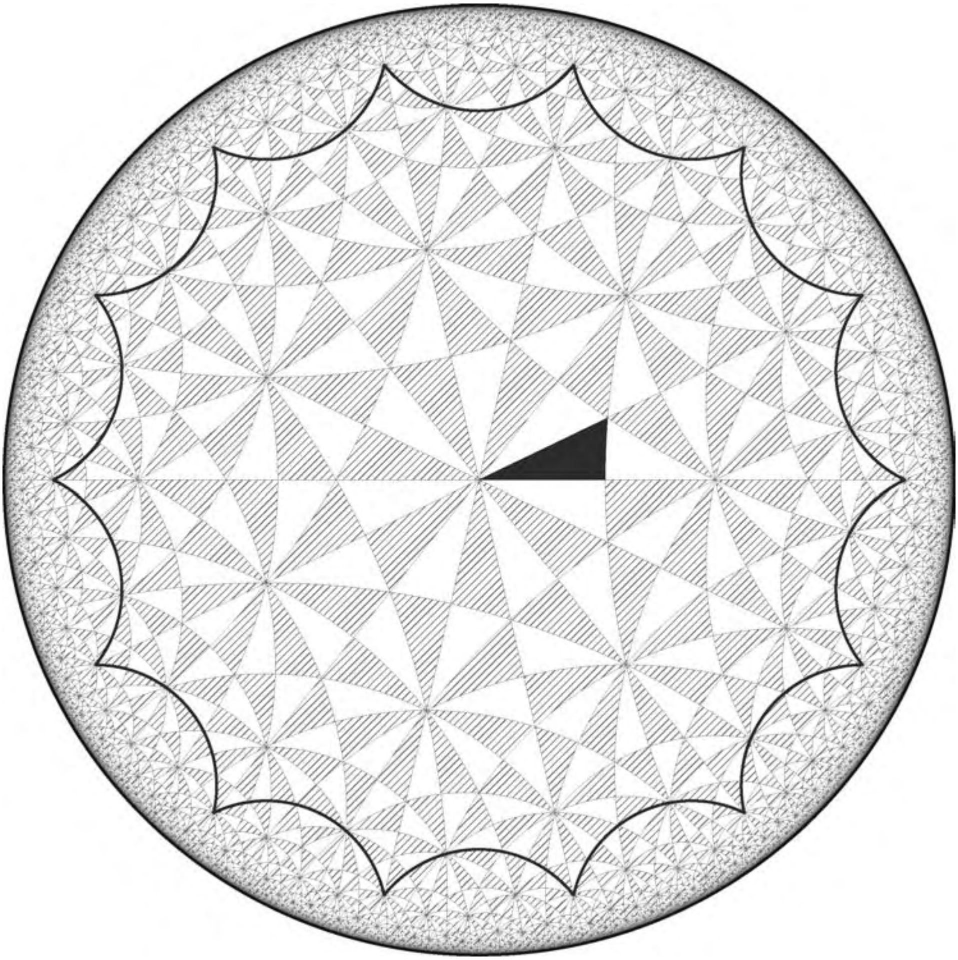


Fig. 2.6. Tessellation by the (2,3,7) group. Two adjacent triangles form a fundamental polygon. A single triangle is a fundamental polygon for the reflection group. A fundamental polygon for the subgroup associated with the Klein surface (page 72) is also indicated.

2-6. *A crash course on tori.* Take $\omega_1, \omega_2 \in \mathbb{C}$ with $\mathrm{Im}(\omega_2/\omega_1) > 0$. Consider the rank two parabolic group

$$G = \langle z \mapsto z + \omega_1, z \mapsto z + \omega_2 \rangle,$$

associated with the lattice in $\mathbb{C}$ of the points $\{m\omega_1 + n\omega_2\}$, $m, n \in \mathbb{Z}$. The parallelogram with vertices $(0, \omega_1, \omega_2, \omega_1 + \omega_2)$ is a fundamental parallelogram: its G-orbit covers $\mathbb{C}$ without overlap. The quotient $T = \mathbb{C}/G$ is a torus. The euclidean metric in $\mathbb{C}$ projects to T and the sides $[0, \omega_1]$, $[0, \omega_2]$ project to a pair of simple loops which cross each other only at the projection of 0.

Set $\omega_2' = a\omega_2 + b\omega_1$, $\omega_1' = c\omega_2 + d\omega_1$, where a, b, c, d are integers satisfying $ad - bc = 1$ (so that ω_1, ω_2 likewise can be expressed as an integral combination of

$\omega_1', \omega_2')$. Thus $\{\omega_1', \omega_2'\}$ is a new basis of the lattice. Every change of basis arises in this manner.

The only conformal mappings of $\mathbb{C}$ are the affine mappings $z \mapsto az + b$. An affine mapping takes one lattice (ω_1, ω_2) to another; it projects to a conformal mapping of one quotient torus to the other, sending one pair of loops to the other. Since we do not want to distinguish two lattices so related, we can normalize by focusing instead on ratios $\tau = \omega_2/\omega_1$, with $\operatorname{Im}\tau \neq 0$, and lattices $\{m + ni\}$.

In this convention $\tau' = \omega_2'/\omega_1'$ determines the same lattice if and only if there exists a normalized Möbius transformation with integer entries a, b, c, d such that

$$\tau' = A(\tau) = \frac{a\tau + b}{c\tau + d}, \quad \tau = \frac{\omega_2}{\omega_1}.$$

Such a transformation A is called a *modular transformation*. It sends each torus to a conformally equivalent one: Instead of the parallelogram $(1, \tau)$ there is a new fundamental parallelogram $(1, \tau')$ for the same lattice. The group of modular transformations is of course the modular group of Exercise 2-5.

We can subject τ to the following additional normalization: $\operatorname{Im}\tau > 0$, $|\tau| \geq 1$, $-\frac{1}{2} < \operatorname{Re}\tau \leq \frac{1}{2}$ but $\operatorname{Re}\tau > 0$ if $|\tau| = 1$ [Ahlfors 1978, §6.2.3]. This choice uniquely determines τ amidst its orbit under the modular group (Exercise 2-5).

The space $\mathfrak{T}$ of all tori can thus be taken to be the upper half-plane $\{\tau : \operatorname{Im}\tau > 0\}$. Two points τ, τ' represent conformally equivalent tori if and only if they differ by a modular transformation. The modular transformations are of course isometries of $\mathfrak{T}$.

A fundamental parallelogram "rolls up" to give the quotient torus T. The specific choice of fundamental parallelogram gives a marking of the torus, in that the two pairs of edges project to an specific ordered pair of simple loops on the quotient torus. (The two loops are geodesics in the euclidean metric of the torus and cross each other exactly once.) Different choices for the fundamental parallelogram correspond to different choices for this pair of simple loops and different choices in the orbit of the initial τ under the modular group.

The change of marking arising from changing τ to τ' is induced by a conformal automorphism of the underlying torus if and only if $\tau' = \tau$ is a fixed point of A, which is then necessarily elliptic. This can happen only for special lattices — special values of τ, namely for $\tau = i$ (square), $\tau = e^{2\pi i/3}$, or $\tau = e^{2\pi i}/6$. The modular transformation that fixes the point $\tau = i$ is $\tau' = -1/\tau$.

There is a continuous group that maps every torus unto itself, without fixed points. There is a unique element that maps a given point to any other. This group is the projection to the torus of the group $\{z \mapsto z + c : c \in \mathbb{C}\}$ of translations of $\mathbb{C}$. By fixing say $z = 0$ as a lattice point, we prevent this group from acting on the quotients.

In addition the map $z \mapsto -z$ projects to every torus T. It becomes a conformal automorphism of order two with exactly four fixed points. The quotient $T/\langle z \mapsto -z\rangle$ is conformally equivalent to $\mathbb{C} \cup \infty$. and T is branched of order two over four distinct points. Apart from the group of translations above, this is the only affine map that induces a conformal automorphism of all tori. Since it has order two, it is called the *hyperelliptic involution*. All closed surfaces of genus 2 also have such an conformal involution which by necessity has six fixed points (see Exercise 2-13), but relatively few closed Riemann surfaces of each $g > 2$ support such an automorphism.

The euclidean line segment from 0 to $p + q\tau$, with (p, q) relatively prime integers, projects to a simple loop, and conversely every simple loop on T from the projection of 0 is determined in such a fashion. In other words, there is a one-to-one correspondence between rational numbers q/p and unoriented simple loops from a given point $0 \in T$. The fraction $0/1$ corresponds to α and $1/0$ corresponds to β.

Actually it is quite artificial to choose the base point 0. Given a slope p/q with p, q relatively prime, consider the family of all parallel euclidean lines with this slope. The projection to T is a family of parallel, mutually disjoint simple loops (geodesics in the euclidean metric) that fills up T. Conversely, each simple loop is freely homotopic to such a family. In short, given any torus T and a basis α, β the set of free homotopy or homology classes of simple closed curves on T is in one-to-one correspondence with the rational numbers.

Let T be the square torus $\tau = i$ and $\alpha \in T$ be the simple geodesic loop coming from a line with slope p/q. If (p, q) are relatively prime, which of course we will always assume, there are relatively prime positive integers (r, s) for which $ps - qr = \pm 1$. Choose $\beta \in T$ to be the simple loop coming from a line with slope r/s. Show that the number of times that α crosses β on T is exactly

$$i(\alpha, \beta) = \left|\det\left(\begin{smallmatrix} p & r \\ q & s \end{smallmatrix}\right)\right|.$$

Here $i(\alpha, \beta)$ is called the *geometric intersection number* of the loops α, β. It is the least number of intersections any pair of curves α' in the free homotopy class of α and β' in the free homotopy class of β can have. Therefore if $h : T \to T$ is a homeomorphism, $i(\alpha, \beta) = i(f(\alpha), f(\beta))$.

Let p, q be relatively prime integers, $0 < p < q$ and consider the line $L : y = (p/q)x$ in the (x, y)-plane. It projects to a simple loop α on the quotient torus T. On T, there is a shortest distance d from one side of α to the other side. Remembering that there is a pair of integers (p', q') with $pp' - qq' = \pm 1$, show that

$$d = \frac{1}{\sqrt{p^2 + q^2}}.$$

Hint: minimize the distance of the lattice point $m + ni$ to the line L. Show that the line $L' : y = (p/q)x + 1/2q$ is as close as possible to L while projecting to the simple geodesic on T parallel to α and halfway between its two sides.

What happens if $\tau \to p/q$, $\tau \in A(F)$, for some $A \in \mathrm{Mod}$? The torus becomes pinched along its (p, q) curve. What happens if $\tau \to \zeta \in \mathbb{R}$ with ζ irrational? Then τ runs through a sequence of polygons $\{A_j(F)\}$ with $A_j \in \mathrm{Mod}$ and $\lim A_j = \zeta$. The sequence of tori, which may be taken to be conformally all the same, collapse to the lines of slope ζ, which is the “ending lamination” for the sequence. We will formally study this idea in a more general context in Chapter 5.

Again work with the square torus T. It is known that any (orientation preserving) automorphism $\neq$ id of T is homotopic to the projection of an affine map f described in terms of column vectors by

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

where a, b, c, d are integers satisfying $ad - bc = 1$. Let λ, λ^{-1} denote the eigenvalues of the coefficient matrix (which therefore is conjugate to $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{smallmatrix}\right)$). If $\lambda > 1$ show that $\lim_{n\to\infty} i(f^n(\alpha), \beta)) = \infty$. Such a map on a torus, which is area preserving, is called an *Anosov mapping*. It preserves two lines through $z = 0$, ℓ_a and ℓ_r, both at irrational angles with respect to π. It stretches one by a factor of λ and compresses the other by a factor λ^{-1}. The projection of each line to T is a geodesic of infinite length, never intersecting itself and dense on T. In contrast, if $|\lambda| = 1$, $\lambda \neq \pm 1$, then f has finite order ($f^n = \text{id}$ for some n). If $\lambda = \pm 1$, f satisfies $f(\alpha) = \alpha$ for some simple geodesic loop α on T. The map f is what we will later call a Dehn twist.

An affine map $A(z) = \alpha z + \beta$ maps the lattice $\Lambda = (1, \tau)$ onto itself if and only if $\alpha \in \Lambda$ and $\alpha\tau \in \Lambda$. Of course this is satisfied if $\alpha \in \mathbb{Z}$. For particular values of τ it can be that $\alpha \notin \mathbb{Z}$. For these cases show that τ and likewise α satisfy quadratic equations with integer coefficients.

The jacobian of the mapping is $|\alpha|^2$. That means the fundamental parallelogram $P = (1, \tau)$ is sent to a parallelogram of $|\alpha|^2$-times the area of P; the image covers the torus T $|\alpha|^2$-times, which is necessarily an integer. Alternately P can be subdivided into $|\alpha|^2$ subparallelograms so that the image of each covers the torus once. The induced (analytic) mapping A_* of the torus onto itself has degree $|\alpha|^2$.

How many fixed points does A_* have? How many distinct solutions mod(Λ) does $(\alpha - 1)z = 0$ have? Well $|\alpha - 1|^2$ is the jacobian of the map $z \mapsto (\alpha - 1)z$ so it induces a map of the torus that covers itself $|\alpha - 1|^2$-times, and by necessity this too is an integer. So A_* has exactly $|\alpha - 1|^2$-fixed points. Likewise the n-th iterate $A^n(z) = \alpha^n z + \beta_n$ has $|\alpha^n - 1|^2$ fixed points. These become dense in T as $n \to \infty$, if $|\alpha| \neq 1$, that is, if A_* is a $n \geq 2$ to 1 analytic mapping of T onto itself.

For example, if $\alpha = i$ then we must have $\tau = i$. The map A_* has degree one and two fixed points, namely the points on T corresponding to $z = 0$ and $z = (1 + i)/2$.

2-7. Suppose G is a noncyclic discrete group all of whose elements fix ∞. Suppose the subgroup G_0 of parabolic transformations is a rank two parabolic group. Then G is the extension of G_0 by at least one of the following: an element of order two (possible in all cases), an element of order four (possible only if a fundamental parallelogram P for G_0 is a square), or an element of order three or six (P is a rhombus with a $\pi/3$ vertex angle).

2-8. Suppose G is discrete. Take all the matrices corresponding to G and replace them by their complex conjugates. Show that the resulting group G' is also discrete and $\Omega(G') = J(\Omega(G))$, $\Lambda(G') = J\Lambda(G)$, where $J(z) = \bar{z}$. In fact $G' = JGJ$.

If two elements A, B in a discrete group without elliptics satisfy $A^p = BA^qB^{-1}$, $p, q \neq 0$, then B preserves the fixed point set of A. Furthermore $p = q$ and either A, B lie in a one or two generator parabolic subgroup or both are powers of a loxodromic element C.

2-9. *Punctured tori and the Farey sequence.* A once punctured torus has a hyperbolic metric but a torus has only a euclidean metric. Yet topologically and analytically there is a close relationship as a once punctured torus corresponds to a choice of basepoint on the torus.

A group G representing a once-punctured torus in $\mathbb{H}^2$ is given by two loxodromic generators X, Y without common fixed point and with parabolic commutator $K = XYX^{-1}Y^{-1}$. Prove that XY cannot fix ∞. *Hint:* If XY fixes ∞ show that YX also fixes ∞ and must be parabolic. Then show that X and Y must fix ∞ (hint: $YX = X^{-1}(XY)X$).

A once-punctured torus is called *square* if it has two simple geodesics α, β which have the same length and cross each other exactly once. Show that these curves are the *systoles* for the surface — the geodesics that have the minimum length among all geodesics on the surface. Automatically $\alpha\beta\alpha^{-1}\beta^{-1}$ is freely homotopic to a simple loop that is retractable to the puncture.

Show that the matrices

$$A = \begin{pmatrix} -1+\sqrt{2} & 0 \\ 0 & 1+\sqrt{2} \end{pmatrix}, \quad B = \begin{pmatrix} \sqrt{2} & 1+\sqrt{2} \\ -1+\sqrt{2} & \sqrt{2} \end{pmatrix}$$

determine a square torus. Find a fundamental polygon in UHP.

Find the matrix generators for the once-punctured torus that corresponds to a regular euclidean hexagon with puncture at the center. What are its symmetries?

The *Farey sequence* $\mathcal{F}$ is very useful in studying once punctured tori. It is based on the modular diagram. The Farey sequence is the orbit of the boundary ∂F_2 of the fundamental polygon for M_2 presented in Exercise 2-5. What is interesting is its labeling.

Note that the orbit of the ideal vertex $\{\infty\}$ under Mod is the set of all rational numbers $\mathbb{Q}$ (plus ∞). Prove:

(i) The rational numbers m/n and p/q are ideal vertices of the tile $g(F_2)$, $g \in M_2$ if and only if $mq - np = \pm 1$.

(ii) The ideal vertex x/y of a tile $g(F_2)$ separates the other two ideal vertices $m/n < p/q$ of $g(F_2)$ if and only if

$$\frac{x}{y} = \frac{m+p}{n+q}.$$

Each irrational number ζ is the limit of a nested sequence of geodesics in the orbit of ∂F_2. For example the sequence

$$\ldots, [p/q, m/n], [p/q, (p+m)/(q+n)], [(2p+m)/(2q+n)], \ldots$$

The sequence can be described in terms of a left-right pattern in the edges of the tessellation $\{M_2(F_2)\}$. There is a wonderful description of the Farey sequence in [Mumford et al. 2002].

What we want to point out here is that the geodesic with endpoints m/n, p/q represents a pair of simple closed geodesics α, β on the punctured torus with the property that α crosses β exactly once. Here α has slope m/n and β has slope p/q — see Exercise 2-6.

2-10. Here are some volumes and areas to compute. G is a nonelementary kleinian group.

(i) Consider the group $P_2 = \langle z \mapsto z+1, z \mapsto z+\tau\rangle$, where $\operatorname{Im}\tau > 0$. The horoball $\mathcal{H} = \{(z,t) \in \mathbb{H}^3 : t > a > 0\}$ is invariant under P_2. Assume that its projection $C(\tau)$ to $\mathcal{M}(G)$ is embedded; this is a solid cusp torus. Using the hyperbolic volume form $dV = dx\,dy\,dt/t^3$ show that for $\theta = \arg\tau$,

$$\operatorname{Vol}(C(\tau)) = \frac{|\tau|\sin\theta}{2a^2}, \quad \operatorname{Area}(\partial C(\tau)) = \frac{|\tau|\sin\theta}{a^2}, \quad \frac{\operatorname{Vol}(C(\tau))}{\operatorname{Area}(\partial C(\tau))} = \frac{1}{2}.$$

For the universal horoball, $a = 1$. By Exercise 2-5 a fundamental parallelogram for P_2 can be chosen so that $|\tau| \ge 1$ and $\frac{\pi}{3} \le \theta \le \frac{2\pi}{3}$. Thus the volume of the solid cusp torus is not less than $\frac{\sqrt{3}}{4}$.

(ii) Likewise assume that the equidistant tube $T(r)$ of hyperbolic radius r about a closed geodesic of length L is embedded in $\mathcal{M}(G)$. Show by working in the upper half-space model that

$$\operatorname{Vol} T(r) = \pi L \sinh^2 r, \quad \operatorname{Area} \partial T(r) = 2\pi L \sinh r \cosh r,$$
$$\frac{\operatorname{Vol} T(r)}{\operatorname{Area} \partial T(r)} = \tfrac{1}{2}\tanh r \nearrow \tfrac{1}{2} \quad \text{as } r \nearrow \infty,$$
$$\operatorname{Vol} T(r) = \tfrac{1}{2}\operatorname{Area} \partial T(r) + \tfrac{1}{2}\pi L(e^{-2r} - 1).$$

(iii) In the case that $\mathcal{M}(G)$ has finite volume, borrowing terminology from Section 3.4, deduce that

$$\operatorname{Vol}(\mathcal{M}(G)) \le \operatorname{Vol}(\mathcal{M}(G)^{\text{thick}}) + \tfrac{1}{2}\operatorname{Area}(\partial\mathcal{M}(G)^{\text{thick}}). \tag{2.6}$$

In fact $\mathcal{M}(G)^{\text{thick}}$ is a compact submanifold whose complement consists of a finite number of cusp tori and tubes about short geodesics.

(iv) Suppose $\mathcal{M}(G)$ is a closed manifold (compact, without boundary). Show that there is a shortest closed geodesic γ. If γ has length L, show that it has an embedded tubular neighborhood of radius $L/4$. *Hint:* expand the tubular neighborhood until at radius r it first touches itself at a point p. The two orthogonals of length r from p to γ, with a segment of γ of length $\le L/2$, form a closed loop. Its length must be $\ge L$.

2-11. A *gaussian integer* is a number of the form $p + iq$, where p, q are integers. The group Γ of normalized matrices whose entries are gaussian integers is called the

Picard group. Show that it is a discrete group but its limit set is the whole sphere. *Hint:* It is generated by the four parabolic transformations,

$$S(z) = z + 1, \quad T(z) = \frac{z}{-z+1}, \quad U(z) = z + i, \quad V(z) = \frac{z}{iz+1}.$$

However note that $T = ASA$, $V = AUA$, where $A(z) = -1/z$. Furthermore, if $B(z) = -z$, $A = TST$ and $B = UAU^{-1}AUA$. Therefore the Picard group is generated by three elements $\langle S, T, A\rangle$. For details, including an explicit fundamental polyhedron see [Wielenberg 1978].

The same reference discusses a variety of subgroups of finite index without elliptic elements of Γ. These interesting subgroups give rise to quotient spaces which are homeomorphic to a variety of knot and link complements, including the Borromean rings.

2-12. *Equality in Jørgensen's inequality.* There are continuous families of geometrically finite groups (Section 3.6) as well as uncountably many nonconjugate, nonelementary, geometrically infinite 2-generator discrete groups that give equality in Jørgensen's inequality [Jørgensen et al. 1992]. In these cases A must be elliptic or parabolic [Jørgensen 1976]. The examples are typically extensions of the modular or other triangle groups. The Picard group is one such extreme group.

However in the class of fuchsian groups, only the triangle groups G (Exercise 2-5) with signature $(2, 3, q)$ with $7 \le q \le \infty$ give equality [Jørgensen and Kiikka 1975]. Confirm that this is the case for the first few examples.

Prove that if $\langle A, B\rangle$ gives equality, then $\langle A, B_1 = BAB^{-1}\rangle$ is also a nonelementary, discrete group which gives equality in Jørgensen's inequality. *Hint:* You will need the identity,

$$\operatorname{tr}(AB_1A^{-1}B_1{}^{-1}) = [\operatorname{tr}(ABA^{-1}B^{-1}) - 2][\operatorname{tr}(ABA^{-1}B^{-1}) - \operatorname{tr}^2(A) + 2],$$

and its consequence,

$$\left|\operatorname{tr}(AB_1A^{-1}B_1{}^{-1}) - 2\right| \le \left|\operatorname{tr}(ABA^{-1}B^{-1}) - 2\right|.$$

See also Lemma 1.5.6.

Using the property proved in the preceding paragraph, prove that if $\langle A, B\rangle$ gives equality, then A is either elliptic of order at least 7 or is parabolic.

Prove that if two Möbius transformations A, B with equal traces generate a nonelementary discrete group, then [Jørgensen 1981]

$$\left|\operatorname{tr}(ABA^{-1}B^{-1}) - 2\right| > \tfrac{1}{8}.$$

2-13. *Genus-two surfaces.* In the disk model find a regular hyperbolic octagon with vertex angles $\pi/4$. *Hint:* start with a tiny octagon centered at $z = 0$. It is nearly a regular euclidean octagon with vertex angles $3\pi/4$. Now increase the distance of the vertices from the origin; the vertex angles strictly decrease to zero, as they become ideal vertices.

Next, in the positive direction label the sides of the octagon $\mathcal{P}$ as $a_1, b_1, a_1^{-1}, b_1^{-1}, a_2, b_2, a_2^{-1}, b_2^{-1}$, as in Figure 2.1 on p. 73. Find an isometry A_i that maps a_i onto a_i^{-1} but sends the positive direction along a_i to the negative direction along a_i^{-1}, that is $A_i(\mathcal{P})$ is adjacent to $\mathcal{P}$ along the exterior side of a_i^{-1}. Similarly find B_i. By starting with a vertex p of side a_1, show that the commutator product (starting from the right) satisfies $B_2^{-1}A_2^{-1}B_2A_2B_1^{-1}A_1^{-1}B_1A_1 = \mathrm{id}$. This is called a *vertex relation*; it says that in the orbit of $\mathcal{P}$, successive images of $\mathcal{P}$ are arranged in the indicated cyclic order about a vertex. Show that the orbit of $\mathcal{P}$ under the group $\langle A_1, A_2, B_1, B_2\rangle$ covers $\mathbb{H}^2$ without overlap. Show that the quotient surface R is a genus 2 surface, the eight vertices project to a single point O, each side projects to a simple loop from O, the loops are mutually disjoint except at O, and they bound a simply connected region Δ on R. The vertex relation is a consequence of the fact that if you make a complete circuit of $\partial\Delta$, the resulting loop is contractible to a point.

Does the rotation by $\pi/4$ of the octagon induce a conformal mapping of R onto itself? How about rotation by π? How about reflection about the geodesic between two opposite vertices. What fixed points on R do the induced mappings have?

Instead of the above pattern, label the edges in the sequence $a_1, b_1, a_2, b_2, a_1^{-1}, b_1^{-1}, a_2^{-1}, b_2^{-1}$ and repeat the process, pairing the opposite sides of $\mathcal{P}$ as before. Find the vertex relation. The quotient gives another surface of genus 2 but for which the simple loops are arranged in a different pattern. Consider the rotation of $\mathcal{P}$ by π. Confirm that on the quotient, it maps each simple loop to its inverse, and has exactly six fixed points. This involution is called the *hyperelliptic involution* J; every closed surface of genus two has one. The quotient $R/\langle J\rangle$ is a sphere with six branch values of order two, that is, R is a two-sheeted cover of $\mathbb{S}^2$, branched over six points.

Conversely, given six distinct points in $\mathbb{S}^2$, the two-sheeted cover branched over the six points is a closed surface of genus two. The covering surface is determined by a normal subgroup of index two in the fundamental group of the 6-punctured sphere; can you find the subgroup?

Show using fuchsian groups that every closed genus-2 surface has a hyperelliptic involution. Hint (following Jørgensen): Express the basic relation

$$ABA^{-1}B^{-1}CDC^{-1}D^{-1} = 1$$

as $ABA^{-1}B^{-1} = DCD^{-1}C^{-1}$. Set $U = A^{-1}B^{-1}C$ and $V = DC^{-1}BA$. Conclude that $ABUV = VUBA$. Think then of a double bagel with A, B simple loops around the holes and U, V through the holes. Skewer the double bagel, puncturing it at 6 points, and then rotate by 180°. The involution is given by $A \to A^{-1}$, $B \to B^{-1}$, $C \to C^{-1}$, $D \to D^{-1}$. See also the example of Section 7.2.

2-14. *No tangents at loxodromic fixed points.* Show that a limit set $\Lambda(G)$ cannot have a tangent line at a fixed point of a loxodromic $g \in G$ with $\mathrm{tr}(G) \notin \mathbb{R}$ unless $\Lambda(G)$ is Möbius equivalent to a circle. Therefore there are no smooth limit sets except circles and euclidean lines. See also Exercise 3-21.

Outline of proof. [Lehto 1987, Lemma 4.2]. Assume $\Lambda = \Lambda(G)$ is not Möbius equivalent to a circle. Suppose to the contrary that $\mathbb{R}$ is the tangent line to Λ at $z = 0$ and there is a loxodromic $g \in G$ of the form $z \mapsto ke^{i\varphi}z$, $0 < k < 1$, $0 \le \varphi < 2\pi$.

If $\varphi = 0$, find $z \in \Lambda$, $\operatorname{Im} z \neq 0$. Then for all n, $\arg g^n(z) = \arg z \neq 0, \pi$. There is a contradiction as $n \to \infty$. If instead $\varphi = \pi$, upon working with g^2 we likewise get a contradiction.

More generally set $\phi = \min\{\varphi, |\pi - \varphi|, 2\pi - \varphi\}$ so that $0 < \phi \le \pi/2$. Construct the symmetric wedges of angle ϕ centered along $\mathbb{R}$: $V = \{re^{i\theta} : \theta \in (-\phi/2, \phi/2)\}$, $V' = \{re^{i\theta} : \theta \in (\pi - \phi/2, \pi + \phi/2)\}$. Thus if z lies in $V \cup V'$, then $g(z) \notin V \cup V'$. Choose a sufficiently small disk D about 0 so that $\Lambda \cap D \subset (V \cup V') \cap D$. Then choose $z \neq 0$ in $\Lambda \cap D$ sufficiently small so that $g(z) \in D$. But then $g(z) \in \Lambda \cap D$ yet $g(z) \notin V \cup V'$, a contradiction.

2-15. Prove a Möbius transformation of the form $A = \left(\begin{smallmatrix} a & b \\ b & d \end{smallmatrix}\right)$, $ad - b^2 = 1$, satisfies $JAJ = A^{-1}$, where $J(z) = -z$. Conclude that if a discrete group G is generated by elements of this form, then $\Lambda(G)$ is invariant under J.

2-16. *Modeling conformal groups by Möbius groups.* Suppose G is a group of Möbius transformations preserving $\Omega \subset \mathbb{C}$ and $\phi : G \to H$ is an isomorphism onto a group of conformal automorphisms H that map another region Ω' onto itself. Suppose $\Phi : \Omega \to \Omega'$ is a quasiconformal mapping such that $\Phi g \Phi^{-1} = \phi(g)$ for all $g \in G$. Prove that there is a conformal mapping $\Psi : \Omega' \to \Omega^*$ such that $\Psi h \Psi^{-1}$ is a Möbius transformation for all $h \in H$. Thus if the action of H on Ω' can be modeled by the action of a group of Möbius transformations, we can find a region Ω^*, conformally equivalent to Ω', where the group H actually becomes a group of Möbius transformations [Maskit 1968].

Hint: Confirm that the Beltrami differential $\mu = (\partial\Phi/\partial\bar{z}) \big/ (\partial\Phi/\partial z)$ satisfies

$$\mu(gz)\frac{\overline{g'(z)}}{g'(z)} = \mu(z) \text{for all } g \in G.$$

Extend μ to $\mathbb{S}^2$ by setting it equal to zero in the complement of Ω. Then solve the corresponding Beltrami equation on $\mathbb{S}^2$; the solution $\Psi : \Omega \to \Psi(\Omega) = \Omega^*$ is uniquely determined if we require it to fix three prescribed points. Show that $\psi : g \mapsto \Psi g \Psi^{-1}$ is an isomorphism of G to a group of Möbius transformations H^* preserving Ω^*. Show that $\Psi \circ \Phi^{-1} : \Omega' \to \Omega^*$ is a conformal mapping inducing the isomorphism $\psi \circ \phi^{-1} : H \to H^*$.

2-17. *Geometric group theory.* An abstract finitely generated group H can be investigated *geometrically* by analyzing its action on the *Cayley graph* $\Lambda(H)$. The graph is constructed as follows. Select a generating set $\mathcal{G} = \{g_1, g_2, \ldots, g_r\}$. We will assume that if $g_k \in \mathcal{G}$ then also $g_k^{-1} \in \mathcal{G}$; the identity is not put in $\mathcal{G}$. The *vertices* of the graph $\mathcal{H}$ are the distinct elements of H. An (oriented) *edge* is a triple $E = (h_1, h_2; {g_k}^{\varepsilon})$, where $h_1, h_2 \in G$, $h_2 = h_1 {g_k}^{\varepsilon}$, and $\varepsilon = \pm 1$. The initial point of E is h_1, the terminal point is h_2 and the *inverse* of E is $E^{-1} = (h_2, h_1; {g_k}^{-\varepsilon})$. There is a special case when the element g_k has order two: then the edge E can be regarded as unoriented. If O is the vertex that corresponds to id $\in H$, then any word in the designated generators is uniquely represented by a path of oriented edges starting from O. The word is the identity if and only if the corresponding path is a closed loop.

For example the word $g_1 g_2 g_1{}^{-1}$ reading from the left corresponds to the path composed of the successive edges from O:

$$(\mathrm{id}, g_1; g_1), \quad (g_1, g_1 g_2; g_2), \quad (g_1 g_2, g_1 g_2 g_1{}^{-1}; g_1{}^{-1}).$$

If $g_1 g_2 g_1{}^{-1} = \mathrm{id}$ then the third edge is the inverse of (id, g_1; g_1).

Two graphs are called isomorphic if there is a one-one mapping of the vertices and edges of one onto the vertices and edges of the other which preserves orientations.

To embed $\mathcal{H}$ in a particular space, for example in $\mathbb{R}^2$, $\mathbb{H}^2$ or $\mathbb{H}^3$, we have to represent its vertices by distinct points and its edges by smooth arcs or geodesic arcs which are mutually disjoint except for common endpoints.

The Cayley graph $\Lambda(H)$ is connected (why?). If there are no (nontrivial) closed loops the graph is called a *tree*. A closed loop is trivial if it is a succession of edges followed by the succession of edges with the opposite orientations.

Show that for a free group H on one generator with generating set $\{g, g^{-1}\}$, $\Lambda(H)$ is isomorphic to the graph on $\mathbb{R}$ whose vertices are integers and whose edges are directed segments between them. Show that graph of the free group on two generators with a generating set $\mathcal{G}$ of four elements is a tree and draw an embedding in $\mathbb{R}^2$. Use the model of a n-generator Schottky group to embed the Cayley graph of an n-generator free group in $\mathbb{C}$.

A Cayley graph can be made into a metric space by mapping each edge with distinct endpoints onto the unit interval thereby assigning it unit length and proportionally giving smaller lengths to each segment of the edge. If both endpoints of the edge are the same, map it onto the circle of unit length. There is at least one geodesic between any two vertices; its length is the number of edges in a shortest chain in the graph that connects them. This metric determines a topology on the graph.

The *ends* of the graph are defined as follows. For any compact subset K, count the number of unbounded components of $\Gamma \setminus K$. The number of ends of $\Lambda(G)$ is defined to be the supremum of the number of such components over all K. Show that the number of ends does not depend on the generating set. It is known that a Cayley graph has either zero, one, or an infinity of ends.

The Cayley graph allows "visualization" of the group H, generalizing the following classical representations: To construct the Cayley graph of a fuchsian group or kleinian group we can use the tiling by a fundamental polygon or polyhedron centered at a point O, which is not an elliptic fixed point (Section 3.5). Use the generating set determined by the face pairing transformations. Show that the Cayley graph is represented by drawing geodesic segments from O the successive points in the orbit of O under words in the face pairing transformations. The graph then appears as "dual" to the tiling by the orbit of the fundamental region and combinatorially reflects that tiling — see Section 3.4. As the graph gets closer to the boundary, more and more it looks like hyperbolic space itself, especially if you are very farsighted an cannot see the edges clearly. Analogy with this concrete situation often inspires the intuition for finding "geometry" in abstract Cayley graphs and looking for a "sphere at infinity". The Cayley graph has turned out to be a powerful tool to study particular classes of abstract groups.

The group H acts on its Cayley graph. Each $f \in H$ sends a vertex v to fv and an edge $E = (h_1, h_2; {g_k}^{\varepsilon})$ to $fE = (fh_1, fh_2; {g_k}^{\varepsilon} f)$. The group action is an isometry in the path metric. In the classical cases at least, we can find a connected finite subgraph that serves as a fundamental set for the action.

Even the family of finitely presented groups is too general to deal with; for example, the question of deciding whether a given element of the group is the identity is known to be undecidable; such groups do not seem amenable to a geometric approach. It was Thurston's work that brought modern combinatorial group theory back to its historic, geometrical roots. In his famous 1987 paper, Gromov presented a condition on the groups that would make possible an effective geometric theory. His definition models a certain property of isometry groups of hyperbolic space.

An abstract, infinite, finitely generated group H is said to be *negatively curved* or *δ-hyperbolic* or *word-hyperbolic* or simply *hyperbolic* if it has the following property: There exists a universal constant $\delta > 0$ such that for *any* geodesic triangle* Δ in the Cayley graph $\Lambda(H)$, a point on one side of Δ lies within distance δ of the union of the other two sides, however long the sides are. This property is called *δ-thinness* or the *Rips thin triangle property*. We know it holds for hyperbolic triangles. The thinness property is independent on the chosen generating set of H. (What does "zero thinness" mean?)

The condition of thinness is a global condition that suggests that the Cayley graph in the large "looks" like hyperbolic space. The theory was first outlined in [Gromov 1987]. It now occupies a large place in combinatorial group theory (see the wonderful expositions [Cannon et al. 1997], [Cannon 2002], which has many explicit examples, [Cannon 1991] or [Ohshika 2002]). The study of hyperbolic groups involves a kind of discrete hyperbolic geometry in the large. A word-hyperbolic group has an "space at infinity", which for kleinian groups is $\mathbb{S}^2$, and which serves a fundamental role in analyzing the group. Within the theory a major question is how to determine whether a hyperbolic group is isomorphic to a kleinian group, more specifically ,

Cannon's Conjecture. A hyperbolic group with boundary homeomorphic to $\mathbb{S}^2$ is isomorphic to a kleinian group representing a closed manifold.

The boundary ∂H of a hyperbolic group H is defined as equivalence classes of geodesic rays with a "natural topology" in terms of the action on the rays by H (as is to be expected from knowledge of hyperbolic space).

According to Gromov, randomly chosen groups are hyperbolic. Measure the complexity of a group given by n generators and a finite number of relations in the generators $\{r_k\}$ by $N = n + \sum_k \text{length}(r_k)$. Let A_N denote the number of groups with

* A geodesic triangle consists of three vertices and three geodesic segments connecting them.

complexity $\leq N$ and let H_N be the number of hyperbolic groups with complexity $\leq N$. It is a theorem [Ol'shanskiĭ 1992] that $\lim_{N\to\infty} H_N/A_N = 1$!

Word hyperbolic groups include all finite groups, finitely generated free groups, fundamental groups of closed hyperbolic n-manifolds. Yet not all groups are hyperbolic, for example, a 2-generator abelian group. Its Cayley graph is a square lattice in $\mathbb{C}$. In turn, hyperbolic groups are a special class of *automatic groups* as in finite state automata. This is the class of groups that can be effectively analyzed by computer, see [Epstein et al. 1992; Ohshika 2002]. The theory originated with a paper of Cannon as distilled by Thurston, and was extensively developed in The Geometry Center (1988–1994). Automatic groups are finitely presented, and have a solvable word problem. An extension to the theory admits the fundamental groups of finite volume hyperbolic manifolds which are not closed. A finitely presented group is hyperbolic if and only if its Cayley graph satisfies a *linear isoparametric inequality*, while if it is automatic it satisfies a *quadratic isoparametric inequality*.

Actually the notion of "hyperbolic" is not restricted to graphs. Any metric space with the property that there is a geodesic between any two points can be considered from the point of view of hyperbolicity — see [Ohshika 2002].

Hyperbolic groups have a certain "negative curvature" while abelian groups have more of a zero curvature (think tori!). Interesting groups may have a negative-like structure yet may also include some special abelian subgroups. An example is the mapping class group, that is, the group of homotopy classes of orientation preserving homeomorphisms of a surface onto itself. To remedy this situation in many cases, Farb [1998] introduced the concept of relatively hyperbolic groups.

Here is the definition in the simplest situation. Suppose H is a finitely generated group and $\Lambda(H)$ is its Cayley graph. Let $G \subset H$ be a finitely generated subgroup. Form a new graph $\widehat{\Lambda}$ as follows. For each $h \in H$ identify all the vertices of $\Lambda(H)$ that correspond to elements lying in the left coset hG.

H is said to be *relatively hyperbolic with respect to* G if $\widehat{\Lambda}$ is Gromov hyperbolic. One can extend this definition to a finite number of finitely generated subgroups G_i. For example if $\mathcal{M}(H)$ has finite volume with one cusp then H is hyperbolic relative to the maximal parabolic subgroup associated with the cusp.

With this new definition, the mapping class group and Teichmüller space itself become relatively hyperbolic [Masur and Minsky 1999].

2-18. *More on Schottky groups.* Prove that a freely generated, purely loxodromic fuchsian group G (which is necessarily a group of the second kind) acting in the upper and lower half-planes, normalized so that ∞ is a limit point, is a classical Schottky group, where the Schottky circles are orthogonal to $\mathbb{R}$. *Hint:* Start with the simplest case that UHP$/G$ is a torus with one boundary component so that G has 2-generators. There are four circles, and the opposite, not adjacent, circles are paired. (If instead the adjacent circles are paired the quotient is a 2-holed disk.) When does the converse hold?

What happens when the 4-Schottky circles form a chain of mutually tangent circles with respect to $\mathbb{S}^2$? (Answer: When the pairing is again opposite there results a once punctured torus and the group becomes fuchsian of the first kind; the fractal dust of a Schottky groups congeals to $\mathbb{R}$.)

Really, in talking about Schottky groups, to a large degree it makes little matter, in describing the construction, if there are tangencies of circles, so long as they are arranged in pairs such that the pairing elements are loxodromic or parabolic sending the exterior of one circle onto the interior of its partner. When a point of tangency is fixed by a parabolic it becomes a puncture, however if the point is not so fixed it does not necessarily become a puncture — see [Gilman and Waterman 2003]. In particular, any finitely generated fuchsian group such that the quotient is a finitely punctured (≥ 1) closed surface is such a limiting case of a circle-Schottky group.

Of course the conditions can be weakened further so that the Schottky circles become Jordan curves. Such will be the case we take general quasiconformal deformations of these fuchsian groups. There are explicit examples given in [Mumford et al. 2002].

Using Ahlfors' Finiteness Theorem (Section 3.1) and Maskit's Planarity Theorem, prove that any finitely generated, free, purely loxodromic kleinian group G with $\Omega(G) \neq \varnothing$ is a Schottky group [Maskit 1967]. A much shorter proof makes use of the convex core to be introduced in Section 3.10: A Schottky group G is characterized by the fact that the convex core of $\mathcal{M}(G)$ is a handlebody (in the case G is fuchsian, we have to take an ε-neighborhood of the convex core). This must be the case for a geometrically finite group that is free and purely loxodromic.

Bringing in the notions of ends and tameness from Section 5.3 for $\mathcal{M}(G)$, together with the Covering Theorem (5.6.1), we can state the following which is particularly interesting when $\Omega(G) = \varnothing$ [Canary 1996, Corollary D]:

Theorem 2.9.1. *Assume G is a finitely generated kleinian group such that $\mathbb{H}^3/G$ has infinite volume. Suppose $H \subset G$ is a finitely generated subgroup of infinite index which is purely loxodromic and free. Then H is a Schottky group.*

Given a set of Schottky circles, consider the group generated by reflections in them. How is this group related to the Schottky group?

Given a set of Schottky circles (mutually disjoint) let U denote their common exterior in $\mathbb{S}^2$. Show that any conformal automorphism of U is the restriction of a Möbius transformation. (*Hint:* Consider the group of reflections in the circles and correspondingly reflect each automorphism g to get a conformal map on the complement of the limit set Λ; the map also extends to be an automorphism of Λ. You will need to use the fact that this set has area zero so that the functions are analytic on this set as well.) Go on to prove that this group of automorphisms is finite — if there are more than two circles. (*Hint:* erect the hyperbolic plane in $\mathbb{H}^3$ on each circle and consider the set of hyperbolic distances between every two of them.) What is the effect of the group of automorphism on the quotient surface $\Omega(G)/G$?

2-19. If g, h generate a discrete group without elliptics, prove that at least one of the four transformations is loxodromic: g, h, gh, gh^{-1}.

2-20. *Rational billiards.* Here is a Riemann surface construction that has been used extensively to study the dynamics of "rational" billiards on a euclidean polygon $P \subset \mathbb{C}$, that is not necessarily convex or even simply connected.

Corresponding to each side e_i of P, place a parallel line e'_i, through the origin. Denote the reflection in e'_i by σ_i. By definition, a *rational billiard table* is one with the property that the group Γ generated by the reflections $\{\sigma_i\}$ is finite; denote the identity by σ_0. If ∂P is connected, this condition is satisfied if and only if each interior vertex angle is a rational fraction of 2π. If ∂P is not connected, this requirement is only a necessary condition for a rational table.

We will also use the notation σ_i to denote the reflection of P in the edge e_i.

Under the assumption that P is a rational table, here is how to glue copies of P together to get a closed Riemann surface.

Let N be the number of distinct elements $\{\gamma_k\}$ of Γ. Take N copies of P each with the labeled edges; denote the copies by $\{P_{\gamma_k}\},\ 1 \leq k \leq N$.

Suppose $\gamma_j = \gamma_i \sigma_m$ for some index m. Then identify the edge e_m of P_{γ_j} with the edge e_m of the reflection $\sigma_m(P_{\gamma_i})$: attach the reflected polygon $\sigma_m(P_{\gamma_i})$ to the polygon P_{γ_j} along the common edge e_m.

Show that with this rule for attachment, an abstract polygon S can be built up from the N tiles. The end result will have no free edges. The vertex angles will be integer multiples of 2π.

Another description of gluing is as follows. Consider the normal subgroup Γ_0 of even index $2M$ in Γ consisting of even numbers of reflections in the lines e'_i. Interpreted as the product of reflections in the edges of P, Γ_0 consists orientation preserving euclidean motions $z \mapsto e^{i\varphi}z + c$. The cosets of $\Gamma_0 \subset \Gamma$ are $\{\Gamma_0 \sigma_j\},\ 0 \leq j \leq m$. Now take the polygon P, and the reflected copies of P, $\sigma_1(P), \sigma_2(P), \ldots$, and glue the edges together using the elements of Γ_0.

The complex structure on S is given by the euclidean coordinates on the polygons, but the vertices must be flattened out by use of $z^{\frac{1}{p}}$ at a vertex with angle sum $2\pi p$.

When P is a rectangle, the group has order 4 and the Riemann surface is a torus; when P is an equilateral triangle, the group has order 10 and the Riemann surface is also a torus. For the theory, see [Masur and Tabachnikov 2002]. The point is that on the surface, a ball starting at a point of P, instead of bouncing off the edges of P runs in a straight line on S, except a billiard path that hits a vertex must end since there is no unique continuation.

2-21. Starting with the finite group of a euclidean polyhedron, can you adjoin other such finite polyhedral groups to obtain a nonelementary kleinian group with singular set forming a specified trivalent graph with the properties specified by Proposition 2.5.2?

2-22. *Homology and simple loops.* Confirm the following folk theorem: Suppose S is a closed, oriented surface of genus $g \geq 1$. Fix a "canonical homology basis" A_i, B_i, $1 \leq i \leq g$. This means $\{A_i, B_i\}$ are simple loops generating the first homology, and A_i crosses B_i once but is disjoint from A_j, B_j for $j \neq i$, that is, each pair (A_i, B_i) corresponds to a "handle". An element γ of the first integral homology group can be written, $\gamma \sim \sum(a_i A_i + b_i B_i)$, where each a_i, b_i is an integer.

Prove that the homology class of γ contains a simple closed curve if and only if the greatest common denominator of $\{a_1 \ldots a_g, b_1 \ldots b_g\}$ is one. For a proof see [Schafer 1976].

2-23. *Belyǐ functions on Riemann surfaces.* A Belyǐ function on the closed Riemann surface R is a meromorphic (rational) function $f : R \to \mathbb{S}^2$ such that each of its critical values is at one of the points $0, 1, \infty$. Here a critical point is a point x where the derivative vanishes, and the corresponding critical value is $f(x)$. Not every closed surface supports such a function; not every Riemann surface can be realized as the branched cover of $\mathbb{S}^2$ with all branch values in $\{0, 1, \infty\}$ (for the topological possibilities see page 63).

Each of the following conditions is necessary and sufficient for R to support a Belyǐ function:

(i) There exists a finite set of points $\{x_i\} \subset R$ such that the Riemann surface $R' = R \setminus \{x_i\}$ is uniformized by a finite index subgroup Γ of the modular group Mod $=$ PSL$(2, \mathbb{Z})$: $R' = \mathbb{H}^2/\Gamma$.

(ii) R' carries a horocycle packing: a collection of mutually disjoint horodisks such that the complement is a union of triangular regions.

(iii) **Belyǐ's Theorem.** R is the Riemann surface determined by an irreducible polynomial equation $P(x, y) = 0$ whose coefficients are algebraic numbers.

A horodisk on R' is the projection of a horodisk at a parabolic fixed point of Γ.

For example for the Riemann surface given by $x^m + y^n = 1$, the projection $f : (x, y) \mapsto x$ has critical values in $\{1, \infty\}$ and so is a Belyǐ function.

In particular Riemann surfaces carrying Belyǐ functions are dense in all Riemann surfaces. Markovic asks: In the moduli space of a closed Riemann surface R, could it be that corresponding to any two Belyǐ surfaces, there a Riemann surface which is an unbranched cover of each?

The proof of the first item is simplest. Let $\pi : \mathbb{H}^2 \to \mathbb{H}^2/M_2$ be the projection to the thrice punctured sphere $S_3 = \mathbb{S}^2 \setminus \{0, 1, \infty\}$ as in Exercise 2-5. If a Belyǐ function f exists on R then $R \setminus f^{-1}\{0, 1, \infty\}$ is a covering surface of S_3 and therefore corresponds to a finite index subgroup of Mod. Conversely suppose $R = \mathbb{H}^2/\Gamma$. Then R is a covering surface of $S = \mathbb{H}^2/$Mod, which is the sphere punctured at ∞ with two branch values. Let $f : R \to S$ be the projection. Let R' denote the result of removing from R inverse images of the branch values on S, and let S' be the sphere punctured at the two branch values and at ∞. Then $f : R' \to S'$ is an unbranched covering. The points on R that we removed are the critical points of f.

For thorough studies of this subject and its relation to oriented trivalent graphs and Grothendieck's "dessins d'enfants", see the beautiful papers [Jones and Singerman 1978; 1996], and also [Brooks 1999].

2-24. *A group is determined by its traces.* Suppose A, B, C are loxodromic without both fixed points in common. Show that the two-generator group $\langle A, B\rangle$ is determined up to conjugacy by the traces of $A, B, C = BA$. Then show that any finitely generated irreducible group (Exercise 2-1) is determined up to conjugacy by the traces of its elements. *Hint:* Normalize A to have fixed points $0, \infty$ and B to have 1 as a fixed point. If the group is generated by $A, B, D, E, \dots$ work in turn with A, D, DA, by temporarily conjugating so that D has fixed point 1, etc. Two parabolics A, B without a common fixed point can be conjugated so that A is the unit translation and B has fixed point at 0. Work out the precise requirements to carry out your argument. Discreteness is not needed for your proof.

2-25. *Subgroups of finite index.* Suppose G is a fuchsian group representing a closed surface $R = \mathbb{H}^2/G$. Show that there are subgroups of finite index k (not necessarily normal subgroups) for any $k \geq 2$. In other words, show that there are k-sheeted, unbranched covering surfaces of R. If H has index k in G, the subgroup $H^* \supset H$ generated by $\{ghg^{-1} : g \in G,\ h \in H\}$ is a normal subgroup of G, of index at most k.

The topological possibilities are described by the Riemann–Hurwitz relation

$$g^* - 1 = k(g - 1),$$

where g^* is the genus of the k-sheeted cover. (*Hint:* If you can topologically find finite-sheeted cover S' of R, you can lift the hyperbolic metric from R to S' to get a conformal cover of R. To find a topological cover, cut R along a simple geodesic, and join two copies of the cut surface by cross identifying along the cuts.) In fact there are only a finite number of index-n subgroups (why?).

Let G_n^* be the *intersection* of all subgroups $\{G_k\}$ of G of index $k \leq n$. Show that G_n^* also has finite index; $\mathbb{H}^2/G_n^*$ is at most a finite-sheeted cover of R and of any other k-sheeted cover of R, where $k \leq n$. (*Hint:* The intersection $G_k \cap G_k'$ has index at most n^2.) Show that $G_{n+1} \subset G_n$.

Define a metric $\rho(\cdot, \cdot)$ on G as follows. Given two elements $A, B \in G$ set

$$\rho(A, B) = \min\left\{\frac{1}{n} : AB^{-1} \text{ lies in a subgroup of index } n\right\}.$$

Thus $\rho(A, B) \leq 1$ and $\rho(A_n, \mathrm{id}) \to 0$ if and only if $A_n \in G_n$ with $n \to \infty$. The completion of G with respect to the metric ρ (called the *profinite completion*) is a compact topological group $\widehat{G}$ homeomorphic to a Cantor set. One then works with the space $\mathbb{D} \times \widehat{G}$. The action of G on this space is $T(z, t) = (Tz, tT^{-1})$, where $T \in G$, $t \in \widehat{G}$. For an exposition and further development of this subject, which leads to an infinite-dimensional Teichmüller-like space called the *universal hyperbolic solenoid*, see [Markovic and Šarić 2004].

2-26. *The groups of regular polyhedra: spherical orbifolds.* Here we will follow the treatment of [Ford 1929]. Let $\mathcal{P}$ be a regular euclidean polyhedron inscribed in the unit sphere $\mathbb{S}^2$. Denote the number of its faces, edges and vertices by F, E, V respectively. By Euler's formula, $F - E + V = 2$. Let ν denote the number of faces at each vertex. When ∂P is projected on $\mathbb{S}^2$ there results a tessellation of $\mathbb{S}^2$ by F regular spherical polygons of vertex angles $2\pi/\nu$. Let μ denote the number of edges bounding each face.

We are interested in the group G of symmetries of $\mathcal{P}$. This is a group of rotations of $\mathbb{S}^2$. There are $2E$ of them: for given an edge $[a_0, b_0]$ and another $[a, b]$, there is a symmetry that sends $[a_0, b_0]$ to $[a, b]$ and to $[b, a]$ in either order.

Here is how to construct a fundamental domain for the action of G on $\mathbb{S}^2$. It will be a spherical triangle (with one exceptional case to be included below).

Choose an edge, to be called the outer edge, and an adjacent face. Join the ends of the edge to the midpoint of the face by two lines we will call inner edges. We have then a euclidean triangle with central angle $2\pi/\mu$.

Project the triangle to $\mathbb{S}^2$, for example by stereographic projection from the plane. We have an spherical triangle σ of central angle $2\pi/mu$ and angle π/ν at the other two vertices.

Let S be the elliptic of order μ that fixes the inner vertex of σ. Locate the midpoint of the outer edge and let T be the elliptic of order two that fixes it. The axes of S and T pass through the center of the ball; S and T rotate $\mathcal{P}$ onto itself.

Prove that $G = \langle S, T \rangle$ with $(T \circ S)^\nu = \mathrm{id}$, and that σ is a fundamental region for its action on $\mathbb{S}^2$.

All the groups in this class are 2-generator groups. The rays from the origin of the ball to the fixed point of S, T, TS on the boundary of σ are pointwise fixed by these three elliptics, and their G-orbit gives the complete set of rotation axes for G.

The possibilities are listed in the following table from [Ford 1929]:

	F	V	E	μ	ν	Order(G)
Tetrahedron	4	4	6	3	3	12
Cube	6	8	12	4	3	24
Octahedron	8	6	12	3	4	24
Dodecahedron	12	20	30	5	3	60
Icosahedron	20	12	30	3	5	60
Dihedron	2	n	n	n	2	$2n$

The dihedron is special in that it has zero volume and two faces which are regular $n \geq 2$ sided polygons inscribed in the equatorial plane.

2-27. *Euclidean orbifolds.* Show that any rank two parabolic group G_0 is a subgroup of a group generated by four elliptics of order two (whose fundamental domain is half a fundamental parallelogram of G_0. This is the $(2, 2, 2, 2)$-group.

Show that the rank two group of the square torus can be generated by two elliptics of order four and one of order two (its fundamental domain is 1/4 of the fundamental square of the rank two parabolic subgroup. This is the $(2, 4, 4)$-group.

Consider the rank two group G_0 whose fundamental parallelogram P is spanned by the vectors 1 and $e^{\pi i/3}$. Show that G_0 is a subgroup of (i) the group generated by two elliptics of order three, fixed points at the two centers of the equilateral triangles T_1, T_2 formed by the diagonal $[1, e^{\pi i/3}]$, and (ii) the group generated by an elliptic of order two with fixed point the midpoint of $[1, e^{\pi i/3}]$, and an elliptic of order three with fixed point the center of T_1. These are the $(3, 3, 3)$-group and $(2, 3, 6)$-group. A fundamental domain of (i) is the equal sided 60° parallelogram with vertices $1, e^{\pi i/3}$ and the centers of T_1, T_2, and of (ii) is the 30° isosceles triangle formed by the diagonal and the center of T_1.

3

Properties of hyperbolic manifolds

In this chapter we gather together basic properties of hyperbolic 3-manifolds. We start with a characterization of their (conformal) boundaries. Then we study the universality of key elements of their internal geometry and the thick/thin decomposition. After that, we study the global structure as revealed by their fundamental polyhedra, and by their convex and compact cores. We introduce the class of manifolds which are essentially compact (geometrically finite); this class is at the core of our studies. Along the way, we describe the class of quasifuchsian groups and digress to take crash courses in 3-manifold surgery and the theory of geodesic laminations. The chapter ends with a description of the rigidity of manifolds of finite volume.

3.1 The Ahlfors Finiteness Theorem

The beginning of the modern theory of hyperbolic manifolds can be pinpointed at the appearance in 1964 of a fundamental result:

Ahlfors Finiteness Theorem [Ahlfors 1964]. *If G is a finitely generated kleinian group, $\partial\mathcal{M}(G) = \Omega(G)/G$ is the union of a finite number of surfaces. Each of them is a closed surface with at most a finite number of punctures and elliptic cone points.*

Punctures arise from rank one parabolic fixed points, and cone (branch) points from elliptic fixed points.

The hyperbolic area formula (Exercise 3-1) implies that a Riemann surface R of genus $g \geq 0$ with $m \geq 0$ cone points of (finite) orders $r_1, \ldots, r_m \geq 2$ and $n \geq 0$ punctures appears as a boundary component of $\partial\mathcal{M}(G)$, for finitely generated, nonelementary G, only if

$$2g + n - 2 + \sum_i \left(1 - \frac{1}{r_i}\right) > 0.$$

In fact the inequality is a necessary and sufficient condition for R to be represented as $R = \mathbb{H}^2/H$ for some fuchsian group H.

Good estimates can be found for the genus and number of punctures of the boundary of kleinian manifolds $\mathcal{M}(G)$ in terms of the number N of generators, and the number of rank one and rank two parabolic conjugacy classes. In particular $\sum g_i \le N$ where g_i is the genus of the i-th component of $\partial\mathcal{M}(G)$. If G is purely loxodromic, then $\partial\mathcal{M}(G)$ has at most $N/2$ components. This calculation is made using the homology considerations of Remark 3.7.2.

The deepest part of Ahlfors' theorem is the assertion that the ideal boundary of a component of $\partial\mathcal{M}(G)$ consists only of punctures — that, in particular, there are no simply connected components. The proof is based on the following idea. The group G, being finitely generated, depends on the finite number of complex parameters in its generating matrices. On the other hand, if a boundary surface R were not of the "finite analytic type" indicated above, then that surface would have an infinite dimensional space of distinct deformations. This results in a contradiction. Recent proofs (see [Kapovich 2001, §4.9] or [Marden 2006], for example) are much easier than Ahlfors' original; in particular, the finiteness assertions on total genus, number of punctures and cone points can be deduced by topological methods. The Ahlfors theorem also follows from the solution of the tameness conjecture (Section 5.4); see Exercise 5-11.

Conjugacy classes of parabolic and elliptic subgroups are not necessarily represented by punctures and cone points in $\partial\mathcal{M}(G)$. Yet these classes are finite too (Exercise 3-15), rounding out Ahlfors' theorem.

3.2 Tubes and horoballs

Consider the axis γ^* of a primitive loxodromic element $g \in G$ (g is a generator of the cyclic loxodromic group fixing γ^*). Suppose first that γ^* is not also the rotation axis of an elliptic element, and that there is no elliptic element (of order two) that interchanges its endpoints. Then γ^* projects to a closed geodesic γ in $\mathcal{M}^{\text{int}}$; the full collection of lifts of γ is the orbit $\{G(\gamma^*)\}$. The length of γ is the length of any segment $[x, gx]$ of γ^*. The loop γ is a simple loop if and only if the orbit of γ^* consists of mutually disjoint geodesics. Conversely, every closed geodesic is the projection of a loxodromic axis γ^*.

Suppose γ^* is taken as the vertical axis from $0 \in \mathbb{C}$ in the upper half-space model. Given r consider the *tubular neighborhood* of radius r about γ^*,

$$N_r(\gamma^*) = \{\vec{x} \in \mathbb{H}^3 : d(\vec{x}, \gamma^*) < r\}.$$

This appears as a euclidean cone with central angle 2θ given by the equation $r = \log(\sec\theta + \tan\theta)$. Alternate expressions of the equation are

$$\tanh r = \sin\theta, \quad \cosh r = \sec\theta, \quad \sinh r = \tan\theta. \tag{3.1}$$

The image of N_r under a Möbius transformation A such that $A(0), A(\infty) \neq \infty$ looks like a banana (Exercise 1-4).

If it is embedded, the projection $N_r(\gamma) = \pi(N_r(\gamma^*)) \subset \mathcal{M}$ is called the *tubular neighborhood of radius* r about γ. The volume and surface area of tubular neighborhoods in $\mathcal{M}(G)$ are presented in Exercise 2-10.

If γ^* is also the axis of rotation of an elliptic element, the projection γ is a closed curve and a cone axis. If there is an elliptic element that interchanges the endpoints of γ^*, then γ^* projects to a finite geodesic segment of length $[x, g(x)]/2$ with endpoints on cone axes, the degenerate case of a closed curve as we go forth and return along the segment.

We now turn to the structure at a parabolic fixed point $\zeta \in \mathbb{S}^2$ of G. Let σ denote any horosphere at ζ. The horosphere has an intrinsic euclidean metric $d_\sigma(\,\cdot\,,\cdot\,)$. There is a parabolic pair $T^{\pm 1} \in G$ for which

$$d_\sigma(x, T^{\pm 1}(x)) \le d_\sigma(x, T_1(x))$$

for all parabolic $T_1 \in \mathrm{Stab}_\zeta$ and all $x \in \sigma$. The same inequality is true for $T^{\pm 1}$ on any horosphere at ζ. We will refer to either $T^{\pm 1}$ as a *least (translation) length parabolic* in Stab_ζ. For the parabolic $x \mapsto x+1$, $d_\sigma(x, x+1) = 1/h$ on the horosphere $\sigma = \{(z, t) : t = h\}$.

We can replace G by a conjugate so that $\zeta = \infty \in \mathbb{S}^2$ and that $z \mapsto z+1$ is a least length parabolic. Suppose for simplicity, ζ is not also fixed by an elliptic element. Then Stab_ζ is either a cyclic parabolic group or a free abelian parabolic group of rank two; we may assume that Stab_ζ is either $\langle z \mapsto z+1\rangle$ or $\langle z \mapsto z+1,\ z \mapsto z+\tau\rangle$, where $\mathrm{Im}\,\tau > 0$ and $|\tau| \ge 1$.

In the former case, the doubly infinite strip $\{z : 0 < \mathrm{Re}\, z \le 1\}$ forms a fundamental domain for its action in $\mathbb{C}$. In the upper half-space model, the slab rising vertically from the strip is a fundamental region for its action in $\mathbb{H}^3$. We see that $\mathbb{C}/\,\mathrm{Stab}_\zeta$ can be viewed as a doubly infinite cylinder; $w = e^{2\pi i z}$ conformally maps it onto $\mathbb{C} \setminus \{0\}$. The quotient $\mathbb{H}^3/\,\mathrm{Stab}_\zeta$ is homeomorphic to $\{z : 0 < |z| < 1\} \times (-\infty, \infty)$, since the quotient of each vertical slice of the slab is conformally equivalent to the punctured disk.

In the latter case, the parallelogram with vertices $\{0, 1, \tau, \tau+1\}$ with two adjacent sides included is a fundamental parallelogram for the action of Stab_ζ on $\mathbb{C}$. The quotient is a torus. The vertical chimney rising from the parallelogram is a fundamental region for the action in the upper half-space model of $\mathbb{H}^3$. The quotient is homeomorphic to $\{z : 0 < |z| < 1\} \times \{z : |z| = 1\}$, where the first factor comes as before from the quotient of vertical slices.

The projection of the horoball $\mathcal{H}_s = \{(z, t) \in \mathbb{H}^3 : t \ge s\}$ may or may not be embedded in $\mathcal{M}(G)$. Once it is embedded for some $t = s$ it will be embedded for all larger values of t. If Stab_ζ is cyclic and $\pi(\mathcal{H}_s)$ is embedded, it is homeomorphic to $\{0 < |z| \le 1\} \times \mathbb{R}$. We refer to this as a *solid cusp tube*, a curtain rod with its axis removed. It has infinite volume and surface area. Its boundary, $\pi(\partial\mathcal{H}_s)$ is called a *cusp cylinder*.

If Stab_ζ has rank two and $\pi(\mathcal{H}_s)$ is embedded, it is homeomorphic to the product $\{0 < |z| \le 1\} \times \mathbb{S}^1$ and is called a *solid cusp torus*. Its boundary is called a *cusp torus*. It has finite volume and surface area by Exercise 2-10. We have defined the solid objects to be closed sets, but we will not always be fastidious in distinguishing one from its interior.

3.3 Universal properties

We will record some important internal properties of the quotient. While the properties are stated for hyperbolic 3-manifolds and orbifolds, they have analogues for hyperbolic surfaces. During the proof we will often rely on the characterization of limits of nonelementary groups to be presented in Theorem 4.1.1. For our use here we will draw from it the following special case.

Lemma 3.3.1. *Suppose* $\{\langle A_n, B_n\rangle\}$ *is a sequence of nonelementary, discrete groups such that* $\lim A_n = A$, $\lim B_n = B$. *Then* $\langle A, B\rangle$ *is also a nonelementary, discrete group. The corresponding conclusion holds as well for a sequence of three generator nonelementary, discrete groups.*

We begin with some notations and definitions.

Given a discrete group G and $x \in \mathbb{H}^3$, for $r > 0$ set

$$\delta_x(r) = \{A \neq \mathrm{id} \in G : d(x, Ax) \leq 2r\}.$$

Define the *injectivity radius* at x as

$$r_x = \mathrm{Inj}(x) = \mathrm{Inj}(G; x) = \inf\{r : \delta_x(r) \neq \varnothing\}.$$

Thus $d(x, Ax) \geq 2r_x$, for any $A \neq \mathrm{id} \in G$; that is, the G-orbit of the ball $\{y : d(x, y) < r_x\}$ has no overlaps. Interpreted at the projection $\pi(x) \in \mathcal{M}(G)$, $\mathrm{Inj}(\pi(x))$ is the radius of the largest embedded open ball centered at $\pi(x)$. On the other hand there exists $A \in \delta_x(r_x)$ such that the points $A^{\pm 1}(x)$ lie on the boundary of the ball of radius $2r_x$ about x.

The injectivity radius is infinite for all points $x \in \mathbb{H}^3$ only when $G = \{\mathrm{id}\}$. As long as x is not on a rotation axis, the radius is positive by the discreteness of G. As $\pi(x)$ approaches a cusp, $\mathrm{Inj}(\pi(x)) \to 0$.

Lemma 3.3.2. *Given* $\delta > 0$ *there exists* $M = M(\delta)$ *such that for any* $x \in \mathbb{H}^3$ *and for any kleinian group* G *with* $r_x = \mathrm{Inj}(G; x) < \delta$, *the set* $\delta_x(r_x)$ *has at most* M *elements.*

Proof. If $A, B \in \delta_x(r_x)$, $A \neq B$, then $d(Ax, Bx) = d(x, A^{-1}Bx) \geq 2r_x$. Therefore the points Ax, Bx on the sphere of radius $2r_x$ about x are of distance $\geq 2r_x$ apart. For fixed $r = r_x$, only finitely many such points are possible. As $r \to 0$, it is approximately the same number as in the euclidean case and this number is uniformly bounded, independent of $x \in \mathbb{H}^3$. □

In the same vein:

Lemma 3.3.3. *Suppose* M *is a hyperbolic surface or manifold with the property that there is a constant* $\delta > 0$ *such that all geodesics have length at least* 2δ *(and there are no punctures or cusps). Then* M *can be covered by embedded* δ*-balls.*

Given $V > 0$, *there exists* $N(V)$ *such that all* M *with area or volume not exceeding* V *can be covered by* $N(V)$ *embedded* δ*-balls.*

Theorem 3.3.4 (Universal properties of kleinian groups). *There exist universal constants in terms of which any nonelementary kleinian group G has the following properties.*

Universal ball. *There exists $\delta > 0$ such that $\mathcal{M}(G)\setminus\{\text{cone axes}\}$ contains an embedded hyperbolic ball of radius δ.*

Universal horoball. *Suppose $\zeta = \infty$ is a parabolic fixed point and $z \mapsto z + 1 \in G$ is a least length parabolic. Then the horoball $\mathcal{H} = \{(z.t) \in \mathbb{H}^3 : t > 1\}$, called the* universal horoball, *satisfies $A(\mathcal{H}) \cap \mathcal{H} = \varnothing$ for all $A \in G$ such that $A(\infty) \neq \infty$. If $A = \left(\begin{smallmatrix} * & * \\ c & * \end{smallmatrix}\right) \in G$, with $c \neq 0$, then $|c| \geq 1$.*

If $\mathcal{H}'$ is the universal horoball at a parabolic fixed point $\zeta' \neq \infty$ of G, then $\mathcal{H}' \cap \mathcal{H} = \varnothing$.

Tubular neighborhoods about short geodesics. *There exist $r > 0$ and $L_0 > 0$ such that in any $\mathcal{M}(G)$:*

(i) The radius r tubular neighborhood about any closed geodesic of length $\leq L_0$ is embedded; any geodesic of length $< L_0$ is simple.

(ii) The r-tubular neighborhoods about different geodesics of length $< L_0$ are mutually disjoint.

(iii) The r-tubular neighborhoods about geodesics of length $< L_0$ do not intersect the universal horoballs.

Universal elementary neighborhood. *There exists $\delta > 0$ such that for any $x \in \mathbb{H}^3$, the subgroup generated by $\{A \in G : d(x, Ax) < 2\delta\}$ is elementary; if the generator A is loxodromic, it represents a simple geodesic.*

Isolated cone (rotation) axes. *There exists $\delta > 0$ such that the distance between any nonintersecting rotation axes in $\mathcal{M}(G)$ is at least δ, except if they have a common endpoint at a rank two cusp, or perhaps if they are axes of order two.*

Proof: The universal horoball.. The universal horoball corresponding to a parabolic T is invariantly defined as that horoball bounded by the horosphere $\sigma = \partial H$ such that for $x \in \partial H$, $d_\sigma(x, T(x)) = 1$ in the intrinsic flat metric on σ. The existence of the universal constant is an immediate consequence of Jørgensen's inequality: If $X = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $Y = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, $ad - bc = 1$, then we find that $\operatorname{tr}(XYX^{-1}Y^{-1}) = 2 + c^2$. Jørgensen's inequality with $A = X$ implies that $|c| \geq 1$, if $c \neq 0$.

The formula for extension of Y to upper half-space $(z, t) \mapsto (z', t')$ says that

$$t' = \frac{t}{|cz + d|^2 + |c|^2 t^2} \leq \frac{t}{|c|^2 t^2} \leq \frac{1}{t}.$$

Hence when $t > 1$, we have $t' < 1$ so that $Y(\mathcal{H}) \cap \mathcal{H} = \varnothing$.

If there is another parabolic S with fixed point $\zeta' \neq \infty$ we may conjugate the group by a translation so that the fixed point of S is moved to $(0, 0)$. The parabolic then has the form $S' : z \mapsto z/(cz + 1)$. We know that $|c| \geq 1$. To see that the universal horoball at 0 is disjoint from the one at ∞ note that $z \mapsto 1/z$ conjugates S' to $z \mapsto z + c$ and its universal horoball there is $(z,t): t > |c| \geq 1$. Returning to S' we see that the boundary of the horoball at 0 meets the vertical axis at $1/|c| \leq 1$.

It is possible that a parabolic fixed point, say $\zeta = \infty$, is also fixed by elliptics $E \in G$. Such elliptics have the form $E(z) = e^{2i\theta}z + a$ and clearly preserve the horoballs at ∞ as well.

Note that if the least length parabolic at ∞ is instead $z \mapsto z + a$, the universal horoball is given by $\{z : \operatorname{Im} z > 1/|a|\}$.

A similar argument applies in $\mathbb{H}^2$. See also Exercise 3-3 on page 164. $\square$

Corollary 3.5.3 presents a version of the horodisk theorem that applies in certain simply connected regions $\Omega \subset \mathbb{C}$ rather than just $\mathbb{H}^2$.

Tubular neighborhoods about short geodesics. If the property (i) does not hold, there are sequences $r_n \to 0$, $L_n \to 0$ and a corresponding sequence of groups G_n and geodesics γ_n such that the radius r_n-tube about γ_n of length $\leq L_n$ is not embedded.

Let ℓ denote the vertical half-line rising from $z = 0$ in the upper half-space model. We may replace each G_n by a conjugate so that γ_n is the projection of ℓ and the corresponding primitive transformation is $A_n : z \mapsto a_n z$, $|a_n| > 1$ where $\log|a_n| \to 0$ is the length of γ_n.

Our hypothesis insures that there is no elliptic of order two in G_n that interchanges the fixed points of A_n. However our proof will still work if we allow elliptics with the same axis as A_n, although the tubular neighborhood will then have a singular axis.

Let C_n denote the euclidean cone about ℓ, which is the radius-r_n tubular neighborhood of ℓ. Let $F_n = \{\vec{x} \in C_n : 1 \leq |\vec{x}| \leq |a_n|\}$ be a fundamental chunk of C_n. There is an element $B_n^* \in G_n$, which does not preserve ℓ, with $B_n^*(C_n) \cap C_n \neq \varnothing$. For some p, q, $B_n = A_n^q B_n^* A_n^p \in G_n$ has the property that $B_n(F_n) \cap F_n \neq \varnothing$. Therefore for some $x_n \in F_n$, $B_n(x_n) \in F_n$. Furthermore $\langle A_n, B_n \rangle$ is not elementary.

After passing to a subsequence if necessary, $\lim A_n = A$ and $\lim B_n = B$ exist. But A, B fix the point $p \in \ell$ with $|p| = 1$ so that $\langle A, B \rangle$ is elementary, a contradiction to Lemma 3.3.1.

Exactly the same proof shows that there cannot be a sequence of groups $\{G_n\}$ in which there are two loxodromics A_n, B_n with translation lengths satisfying $L_n \to 0$, such that the closure of their r_n-tubes intersect, $r_n \to 0$.

To prove (iii), suppose that $A : z \mapsto z+1$ is an element of G and $T \in G$ is loxodromic with translation length L, which is the length of the corresponding geodesic in $\mathcal{M}(G)$. Conjugate the group by a translation so that the fixed points of T are symmetric about $z = 0$. Then a normalized matrix for T has the form $T = \begin{pmatrix} a & bd \\ d/b & a \end{pmatrix}$, $a^2 - d^2 = 1$, $\operatorname{tr}(T) = 2a$, and its fixed points are $\pm b$. Therefore the r-tube about the axis of T will not intersect the horoball $\mathcal{H} = \{(z,t) : t > 1\}$, provided that $|b| \leq e^{-r}$.

We claim that for all sufficiently small $L > 0$, it will be true that $|b| \leq e^{-r}$. For otherwise, there is a sequence of groups G_n containing A and a loxodromic T_n with fixed points $\pm b_n$ symmetric to $z = 0$ such that $\lim L_n = 0$, while $\lim b_n = b^* \geq e^{-r}$. If $b^* = \infty$, then a long segment of the axis of T_n penetrates $\mathcal{H}$ in which case $T_n(\mathcal{H}) \cap \mathcal{H} \neq \varnothing$ in contradiction to the universal horoball property. On the other hand, since $b^* \neq 0$, $\lim T_n = T$ exists with T either elliptic or the identity. This violates Lemma 3.3.1. We conclude that there exists $L' \le L$ for which the r-tube about any geodesic in any $\mathcal{M}(G)$ of length $\le L'$ does not intersect any universal solid cusp torus or cusp tube.

With the proper interpretation, the case that the loxodromic axis is also the rotation axis of an elliptic in G, or is preserved by an element of order two, is included in our analysis. $\square$

For the record we also point out the following interesting inequality [Meyerhoff 1987]. By the universal horoball property $|b| \le |d|$ so that $4|b|^2 \le 4|d|^2 = |\mathrm{tr}^2(T) - 4|$. Now T is conjugate to $\begin{pmatrix} k & 0 \\ 0 & k^{-1} \end{pmatrix}$, where $k = e^{\frac{L}{2}} e^{i\varphi}$, $0 \le \varphi < \pi$, and $\mathrm{tr}^2(T) - 4 = (k - k^{-1})^2$. Consequently

$$|b|^2 \le \sinh^2 \frac{L}{2} + \sin^2 \varphi. \tag{3.2}$$

With the help of a sophisticated computer search, D. Gabai, R. Meyerhoff, and N. Thurston [Gabai et al. 2003] proved that with a few exceptions, if G has no parabolic or elliptic transformations, there is a geodesic in $\mathcal{M}(G)$ with an embedded tubular neighborhood of radius $r = (\log 3)/2$.

Isolated rotation axes. Confirmation of this property runs along the same lines. In a sequence of groups $\{G_n\}$, suppose that $\{d_n,\ d_n > 0\}$ is a sequence with $\lim d_n = 0$, that $E_n \in G_n$ is an elliptic with rotation axis ℓ, and that $F_n \in G_n$ is an elliptic whose axis ℓ_n does not intersect ℓ but comes within distance d_n of ℓ. We may replace G_n by a conjugate so that for some $p_n \in \ell_n$, $\lim p_n = p \in \ell$.

For all large n we may assume that either $\langle E_n, F_n\rangle$ is nonelementary, it is an infinite dihedral group, that is, each of E_n, F_n is of order two and $E_n F_n$ is loxodromic with axis orthogonal to the axes of E_n and F_n, or ℓ, ℓ_n have a common endpoint at a rank two cusp at ∞ with $\langle E_n, F_n\rangle$ a subgroup of $\mathrm{Stab}_n(\infty)$. In the former case for a subsequence, both $\lim E_n = E$, $\lim F_n = F$ are Möbius transformations fixing p. But then $\langle E, F\rangle$ is elementary, again a violation of Lemma 3.3.1. $\square$

The universal elementary neighborhood. Denote by $G_x(r)$ the subgroup of G generated by the set $\delta_x(r)$, that is by the elements A for which $d(x, Ax) \le 2r$. We claim that there exists $r > 0$ such that, for any $x \in \mathbb{H}^3$, and any kleinian group G, the subgroup $G_x(r)$ is elementary. In other words, for the ball $B_x(r)$, the subgroup generated by the elements $\{g\}$ for which $g(\overline{B_x(r)}) \cap \overline{B_x(r)} \neq \varnothing$ is elementary.

For a fixed G and $x \in \mathbb{H}^3$, there must be some $r > 0$ for which $G_x(r)$ is elementary. For as $r = r_n \to 0$, any infinite sequence of distinct elements $A_n \in \delta_x(r_n)$ converges either to an elliptic transformation fixing x or to the identity. No such sequence can exist! So for all sufficiently small r, the set of elements $\delta_x(r)$ is independent of r and either contains no elements or consists of elliptic transformations fixing x.

Now assume that for some $x \in \mathbb{H}^3$, there is no universal elementary neighborhood. Then there is a sequence of kleinian groups $\{G_n\}$ and a sequence $r_n \to 0$ such that $G_{n,x}(r_n)$ is not elementary. On the other hand, for fixed n, $G_{n,x}(\rho)$ is elementary for some ρ with $0 < \rho < r_n$. As ρ increases to r_n, the elementary groups $G_{n,x}(\rho)$ are nested. There is a first number $\tau_n < r_n$ for which $G_{n,x}(\rho) = G_{n,x}(\tau_n)$ for $\tau_n \le \rho < r_n' \le r_n$, and is elementary but $G_{n,x}(r_n')$ not. We may take $r_n = r_n'$.

If $G_{n,x}(\tau_n)$ is finite but not cyclic, there are elements $A_n, B_n \in \delta_x(\tau_n)$ with distinct, yet intersecting, axes of rotation. The set $\delta_x(r_n)$ must contain an element X_n which does not fix the common fixed point of A_n, B_n. Hence $\langle A_n, B_n, X_n\rangle$ is not elementary.

If $G_{n,x}(\tau_n)$ is finite cyclic, let $A_n \in \delta_x(\tau_n)$ be a generator. We can find an $X_n \in \delta_x(r_n)$ that does not fix the axis of A_n. If X_n is elliptic with axis intersecting that of A_n, there must be another element $Y_n \in \delta_x(r_n)$ that does not fix this common point. Thus one of $\langle A_n, X_n\rangle$ and $\langle A_n, B_n, X_n\rangle$ is not elementary.

Next suppose $G_{n,x}(\tau_n)$ is an infinite group that keeps invariant a line $\ell \subset \mathbb{H}^3$. Either $\delta_x(\tau_n)$ contains a loxodromic A_n, or it contains two elliptics A_n, B_n of order two that interchange the endpoints of ℓ. There must be an element $X_n \in \delta_x(r_n)$ which does not leave ℓ invariant. Again, $\langle A_n, X_n\rangle$ or $\langle A_n, X_n, Y_n\rangle$ is not elementary.

Finally, suppose $G_{n,x}(\tau_n)$ fixes a point $\zeta \in \mathbb{S}^2$, but does not fall into one of the previous cases. Then $\delta_x(\tau_n)$ contains a parabolic transformation A_n or two elliptics A_n, B_n such that B_nA_n is parabolic (e.g., $z \mapsto -z$, $z \mapsto -z+1$). The set $\delta_x(r_n)$ contains an element X_n that does not fix ζ. Hence $\langle A_n, X_n\rangle$ or $\langle A_n, B_n, X_n\rangle$ is not elementary.

In all cases we have found a nonelementary two- or three-generator subgroup generated by elements of $\delta_x(r_n)$. As $n \to \infty$, convergent subsequences converge to elements which fix x and are therefore elliptic or the identity. Once again we draw on Lemma 3.3.1 to reach the contradiction.

To complete the argument we claim that, if $A \in G_x(r)$ is loxodromic, it represents a simple geodesic in the quotient. Otherwise the projection of $[x, Ax]$ into $\mathcal{M}(G)$ would contain two simple subloops of shorter length. (The projection $\pi([x, Ax])$ is a closed loop which is a geodesic except for a likely corner at $\pi(x)$.) We could find two other loxodromics A_1, A_2 with different axes, and satisfying $d(x, A_ix) < 2r$. Both elements would be in $G_x(r)$ which then could not be elementary. □

The universal ball. Given an $x \in \mathbb{H}^3$ that is not an elliptic fixed point of G, denote by G_x the subgroup of G generated by the set $\delta_x(r_x)$, where $r_x = \mathrm{Inj}(G; x)$; in other terms, $\delta_x(r_x) = \{A \in G : d(x, Ax) = 2r_x\}$.

Given any nonelementary, discrete group G we will find somewhere in $\mathbb{H}^3$ a ball of radius δ, where δ is the universal elementary neighborhood constant, with the property that there is no overlapping in its G-orbit. To do this we will find a point $x \in \mathbb{H}^3$ for which G_x is not elementary. The universal elementary neighborhood property then assures us that $r_x \ge \delta$, and as a consequence $g(B_x(\delta)) \cap B_x(\delta) = \varnothing$ for all $g \ne id \in G$, for the δ-ball about x. In searching for such a point, we may restrict our attention to groups whose injectivity radii are uniformly bounded above (no elementary discrete groups have this property).

Start with a point $x \in \mathbb{H}^3$ which is not a fixed point. Suppose G_x is elementary. We will find a polygonal line along which the injectivity radius strictly increases until the terminal point y where G_y is nonelementary. To find this line we have to examine various classes of elementary groups separately.

Case 1: The elements of G_x have a single common fixed point $\xi \in \partial\mathbb{H}^3$. We may take $\xi = \infty$ in the upper half-space model. Then G_x is a finite extension of a rank one or rank two parabolic group. For each $A \in \mathrm{Stab}_\infty(G)$, $A \neq \mathrm{id}$, the perpendicular bisectors of the segments $[x, A^{\pm1}x]$ are vertical half-planes. If $A \in \delta_x(r_x)$, they are tangent to the ball D_x of radius r_x about x. For $A \notin \delta_x(r_x)$, the perpendicular bisectors are uniformly bounded away from D_x.

Let the point y move down the vertical line ℓ through x. For a certain open interval near x, $\mathrm{Inj}(y) = r_y$ is determined by the same vertical planes that determine $\mathrm{Inj}(x)$. However r_y is strictly increasing since y is moving closer to $\partial\mathbb{H}^3$, away from ∞. Since we are assuming r_y is uniformly bounded there must be a first point w with the following property. For some $B \in \delta_w(r_w)$, $B \notin \delta_x(r_x)$, the perpendicular bisector of $[w, Bw]$ is tangent to the ball D_w of radius r_w about w. This cannot be a vertical plane since $B \notin \mathrm{Stab}_\infty$. Because $\langle G_x, B\rangle \subset G_w$, the group G_w is not elementary.

Case 2: G_x is a finite group but not a cyclic group nor a $\mathbb{Z}_2$ extension of a cyclic group. We now use the ball model and take the common fixed point of G_x to be the origin. Then G_x is a subgroup of the finite group $\mathrm{Stab}_0 \subset G$ of euclidean rotations. The ball D_x of radius r_x centered at x is inscribed in a convex cone with flat faces and vertex at the origin; its faces are contained in the perpendicular bisecting planes of $[x, A^{\pm1}x]$, for $A \in \delta_x(r_x)$. These are equatorial planes of the ball model $\mathbb{H}^3$.

Let now the point y move away from x along the ray from the origin through x. There is a first point w for which the ball D_y of radius r_y hits a new plane, the perpendicular bisector of some $[w, Bw]$, $B \notin \delta_x(r_x)$. This new plane does not pass through the origin so that $B \notin \mathrm{Stab}_0$. Therefore $\langle G_x, B\rangle \subset G_w$ is not elementary.

Case 3: G_x is cyclic loxodromic or a finite extension of a cyclic loxodromic group. In preparation for the analysis of this case, we take note of the following situation. Suppose, say in the ball model, we have a closed ball B with center on the positive radius of $\mathbb{H}^3$ that does not contain 0. Let P denote the equatorial plane through 0 and orthogonal to the vertical diameter of $\mathbb{H}^3$; it contains the center of B. Let ρ denote the diameter of B that lies in P and is orthogonal to the positive radius. Consider two planes β_1, β_2 tangent to B at the ends of ρ. Compare these two planes with two planes X_1, X_2 containing the vertical diameter of $\mathbb{H}^3$ that are tangent to B, as the pages of a book with spine the vertical diameter. Necessarily the planes X_1, X_2 intersect the planes β_1, β_2 in two lines. Consequently we see that the pages of the book grow apart more quickly than do β_1, β_2, as we head toward $\partial\mathbb{H}^3$ along the positive radius.

Now we can deal with Case 3. Let ℓ denote the axis of a primitive loxodromic element $T \in \delta_x(r_x)$; we may assume that it is the vertical diameter. The ball $B_x(r_x)$ of radius r_x about x is tangent to two planes $\beta_\pm$ orthogonal to ℓ, namely, the perpendicular bisectors of the segments $[x, T^{\pm1}(x)]$.

If $\delta_x(r_x)$ contains a rotation $z \mapsto e^{2i\theta}z$ with axis ℓ then $B_x(r_x)$ is also supported by two planes passing through ℓ and opened at angle 2θ.

If $\delta_x(r_x)$ has an element of order two that interchanges the endpoints of ℓ, then $T = E_2E_1$, where each E_i is such a half-rotation. There are two other planes $\beta'_\pm$ orthogonal to ℓ and tangent to $B_x(R_x)$. One contains the axis of E_1 and the other the axis of its conjugate $E_2E_1E_2$.

Construct the ray ρ orthogonal to ℓ and passing through x. Follow a point $y \in \rho$ as y moves from x towards $\partial\mathbb{H}^3$. The thrust of our initial observation is that $B_y(r_y)$ will eventually no longer be supported by the pages of the open book whose spine is ℓ and G_y will be either cyclic or an infinite dihedral group. As y continues to move out, there will be a first point w such that $\delta_w(r_w)$ contains an element $S \notin \delta_x(r_x)$. The perpendicular bisecting plane of $[w, S(w)]$ cannot contain ℓ, nor can it be orthogonal to ℓ. Consequently $\langle G_w, S\rangle$ is not elementary.

Case 4: G_x is a finite cyclic group or the extension of one by an elliptic of order two that exchanges the fixed points. Once again in the ball model, we can assume the vertical diameter ℓ is the axis of rotation. If G_x is cyclic the ball $B_x(r_x)$ of radius r_x about x is supported by two vertical planes containing ℓ, as between the pages of an open book. If in addition if there is an elliptic of order two $E \in G_x$ exchanging the north and south poles, then $E \in \delta_x(r_x)$ and we can assume in addition that $B_x(r_x)$ is also tangent to the horizontal equatorial plane.

In the latter case, let the point y, as before, move toward $\partial\mathbb{H}^3$ from x along the ray orthogonal to ℓ passing through x. There is a first point w at which the ball $B_w(r_w)$ hits the perpendicular bisecting plane P of $[w, Sw]$ for $S \notin \delta_x(r_x)$. The new plane P does not contain ℓ. If $\langle G_w, S\rangle$ is still elementary, we must return to Cases 1–3.

If G_x is cyclic, follow the same procedure. There is a first point w at which the ball $B_w(r_w)$ hits the perpendicular bisecting plane of $[w, Sw]$ for $S \notin \delta_x(r_x)$. The group $\langle G_w, S\rangle$ is not cyclic but may fall into any of the cases 1–4. $\square$

In the case of fuchsian groups, the tubular neighborhood property is called the *collar lemma*. The first paper on it was by Linda Keen. The sharp statement is this: *On a Riemann surface, the length of any nonsimple closed geodesic must exceed* $4\sinh 1$. If α is a simple closed geodesic of length L, then it has a collar neighborhood of width $2\,\mathrm{arcsinh}((\sinh L/2)^{-1})$ (see Exercise 8-7); for a complete discussion see [Buser 1992, Chapter 4].

There has been much recent work studying tubular neighborhoods, especially as a way of better understanding the volume of manifolds [Meyerhoff 1987; Gabai et al. 2001; Przeworski 2003]. The radius ρ of the tube about a closed geodesic γ can be chosen as a function of the length so that as the length of γ shrinks to 0 (in a sequence of groups), and a primitive loxodromic generator converges to a parabolic transformation, $N(\gamma^*)$ converges to the corresponding universal horoball. Explicit estimates are given in [Meyerhoff 1987].

Historical remarks

The universal horoball property seems to have been discovered by Fatou [1930, p. 159] though, as pointed out by Alan Beardon, his proof was incomplete. Apparently the first complete proof in the literature is in [Shimizu 1963] and in some papers the property is referred to as "Shimizu's lemma".

The universal elementary constant is today usually called the *Margulis constant*. For the case without elliptics, it appears in [Kazhdan and Margulis 1968]. The general case appears in [Wang 1969]. These early results were proved in the context of general Lie groups. Following an entirely different track, in the context of hyperbolic geometry in $\mathbb{H}^3$, the property was discovered in 1973 in discussion with Jørgensen. It was one of a number of universal properties that followed from Jørgensen's inequality. This discovery was motivated by the fuchsian analogue in [Marden 1974d], and independently [Sturm and Shinnar 1974]. Jørgensen's lemma brings the analysis closer to the actual phenomena allowing, in principle, estimates for the optimal value.

In this book we have chosen to call the universal constants by descriptive names.

3.4 The thick/thin decomposition of a manifold

Assume that G has no elliptics; the only elementary subgroups of G are then rank one and two parabolic groups, and cyclic loxodromic groups.

The ε-*thin part* $\mathcal{M}^{\text{thin}}(G)$ of $\mathcal{M}(G)$ is defined as

$$\{x \in \operatorname{Int}\mathcal{M}(G) : \operatorname{Inj}(x) < \varepsilon\}.$$

Here $\operatorname{Int}\mathcal{M}$ denotes the interior of $\mathcal{M}$.

For example, if a geodesic γ has length $s < 2\varepsilon$, then it lies in the ε-thin part (the diameter of the ball of radius ε is 2ε). There is number r such that the r-tube about γ has the property that the length of the shortest curve on its boundary, freely homotopic to γ, has hyperbolic length 2ε. So the r tube is the maximal tube about γ with the property that all points in it have injectivity radius $< \varepsilon$.

Also lying in the ε-thin part are the projection of horoballs corresponding to parabolic rank one or rank two groups P. Assuming $T : z \mapsto z+1$ lies in P, choose the horoball bounded by the horosphere σ with the property that the hyperbolic distance satisfies $d(x, T(x)) = 2\varepsilon$, $x \in \sigma$.*

Let $\varepsilon = \delta$ denote the universal elementary constant. Given $x \in \mathcal{M}(G)$, the set $\{g \in G : d(x, g(x)) < 2\varepsilon\}$ either consists only of the identity, or generates either a cyclic loxodromic or a parabolic subgroup. In the latter cases, x lies in a ε tubular neighborhood about a short geodesic or in a solid cusp tube or cusp torus, the projection of a horoball as described above.

These tubes and cusp tori are necessarily disjoint. For if a point $x \in \mathbb{H}^3$ is common to two of them, there are two elements $g_1, g_2 \in \delta_x(\varepsilon)$ which together do not generate an elementary group, a contradiction to our choice of ε.

We summarize our discussion as follows.

* The intrinsic distance and the hyperbolic distance d between $(-1/2, a)$ and $(1/2, a)$ on the horosphere $\{(z, t) : t = a > 0\}$ are $1/a$ and $d = 2\log((1+\sqrt{4a^2+1})/(2a))$. To have $d = 2\varepsilon$, we must have $a = 1/(2\sinh\varepsilon)$.

Proposition 3.4.1. *Let G be a kleinian group without elliptics. For the universal elementary constant $\varepsilon > 0$, the ε-thin part $\mathcal{M}^{\text{thin}}(G)$ is the union of mutually disjoint components consisting of*:

(i) *The ε-tube about a geodesic of length $< 2\varepsilon$,*
(ii) *The ε-solid cusp tube corresponding to a rank one parabolic subgroup,*
(iii) *The ε-solid cusp torus corresponding to a rank two parabolic subgroup.*

It is shown in [Meyerhoff 1987] that one can choose $\varepsilon = 0.052$.

The complement of $\mathcal{M}^{\text{thin}}$ is called the *thick part* and denoted by $\mathcal{M}^{\text{thick}}$:

$$\mathcal{M}^{\text{thick}} = \{x \in \operatorname{Int} \mathcal{M}(G) : \operatorname{Inj}(x) \geq \varepsilon\}.$$

3.5 Fundamental polyhedra

Fundamental polyhedra provide "concrete" models of the manifolds $\mathcal{M}$. Suppose we are standing at an interior point $\pi(\mathcal{O}) \in \mathcal{M}(G)$ and blow up a balloon. If it keeps growing without ever touching itself, we must be living in $\mathbb{H}^3$ itself. Otherwise at some point the balloon will meet itself. We blow some more, and keep blowing until the balloon fills the whole manifold (ignoring the fact that this may require an infinite volume of air). The balloon will then be the projection of the Dirichlet region centered at $\mathcal{O}$; the faces comprise the balloon surface and form a *spine* for the manifold.

The Dirichlet regions, or Poincaré fundamental polyhedra (Poincaré first used them to study kleinian groups), are constructed as follows. Given a kleinian group G, choose a base point $\mathcal{O} \in \mathbb{H}^3$ which is not a fixed point of G. For each element $g \in G$, $g \neq \text{id}$, construct the hyperbolic plane which is the perpendicular bisector P_g of the geodesic segment $[\mathcal{O}, g^{-1}(\mathcal{O})]$. Denote by H_g the relatively closed half-space which is bounded by P_g and contains $\mathcal{O}$. The labeling is such that $g(P_g) = P_{g^{-1}}$ and $g(H_g)$ is complementary to $H_{g^{-1}}$ but shares with it the bounding plane $P_{g^{-1}}$. This notation is used because it is consistent with what is forced in the construction of the isometric polyhedron where the isometric plane for g in the ball model is the perpendicular bisector of $[0, g^{-1}(0)]$.

The *Dirichlet region* or *Dirichlet fundamental polyhedron* $\mathcal{P}_\mathcal{O}$ with center $\mathcal{O}$ is defined as the closed, convex hyperbolic polyhedron

$$\mathcal{P}_\mathcal{O} = \mathcal{P}_\mathcal{O}(G) = \bigcap_g H_g \subset \mathbb{H}^3.$$

(The use of the word "region", and also "domain", is traditional in this context although the set in question is not open.)

If $h \in G$, $\mathcal{P}_{h(\mathcal{O})} = h(\mathcal{P}_\mathcal{O})$.

The relative boundary of $\mathcal{P}_\mathcal{O}$ is the union of possibly an infinite number of faces $\{f\}$ (a face is a polygonal region in some P_g), edges $\{e\}$ (an edge is a geodesic segment that lies in the boundary of two adjacent faces), and vertices $\{v\}$. At most a finite number of faces, edges, and vertices meet any given compact subset of $\mathbb{H}^3$. Moreover, since $\mathcal{P}_\mathcal{O}$ is convex, its intersection with any hyperbolic plane is connected.

Proposition 3.5.1. $\mathcal{P}_{\mathcal{O}}$ *has the following properties*:

(i) *The faces are arranged in pairs* (σ, σ'). *To each pair corresponds an element* $g \in G$, *called a face pairing transformation*, *such that*

$$g(\sigma) = \sigma' \quad \textit{and} \quad g(\mathcal{P}_{\mathcal{O}}) \cap \mathcal{P}_{\mathcal{O}} = \sigma'.$$

(ii) *If a face pairing transformation is elliptic*, *there is an edge contained in its rotation axis*.

(iii) *To each edge* e *corresponds an edge relation*: $g_1 g_2 \dots g_n = g_e$ *where either* g_e *is elliptic with rotation axis containing* e *and* $g_e^m = \mathrm{id}$ *for some* $m > 1$, *or* $g_e = \mathrm{id}$. *Each* g_i *is a face pairing transformation. The polyhedra*

$$\mathcal{P}_{\mathcal{O}},\ g_1(\mathcal{P}_{\mathcal{O}}),\ g_1 g_2(\mathcal{P}_{\mathcal{O}}), \dots,\ g_1 g_2 \cdots g_e(\mathcal{P}_{\mathcal{O}})$$

are arranged cyclically about e, *each sharing a face with the previous and the succeeding. If* $g_e = \mathrm{id}$ *then* $g_1 g_2 \cdots g_e(\mathcal{P}_{\mathcal{O}}) = \mathcal{P}_{\mathcal{O}}$. *Otherwise the full cycle is completed by applying in succession* $g_e, g_e^2, \dots, g_e^m = \mathrm{id}$ *to the union of the listed polyhedra*.

(iv) *The orbit of* $\mathcal{P}_{\mathcal{O}}$ *under* G *fills* $\mathbb{H}^3$ *without overlap on interiors*.

(v) *The face pairing transformations generate* G; *the edge relations generate the relations in* G.

(vi) $\overline{\mathcal{P}}_{\mathcal{O}} \cap \Omega(G)$ *is a fundamental region for the action of* G *on* $\Omega(G)$. *Here* $\overline{\mathcal{P}}_{\mathcal{O}}$ *denotes the closure of* $\mathcal{P}_{\mathcal{O}}$ *in* $\Omega(G) \cup \mathbb{H}^3$.

(vii) *Let* $B_R(\mathcal{O})$ *be the closed ball of radius* R *centered at* $\mathcal{O}$. *Then the intersection* $\mathcal{P}_{\mathcal{O}} \cap B_R(\mathcal{O})$ *projects to a compact submanifold of* $\mathcal{M}(G)$ *bounded by the projection of* $\mathcal{P}_{\mathcal{O}} \cap \partial B_R(\mathcal{O})$.

Proof. (a) The polyhedron $\mathcal{P}_{\mathcal{O}}$ is characterized by the property that a point $y \in \mathbb{H}^3$ lies in its interior if and only if $d(\mathcal{O}, y) < d(y, h^{-1}(\mathcal{O})) = d(\mathcal{O}, h(y))$ for all $h \neq id \in G$. Thus $\mathrm{Int}(\mathcal{P}_{\mathcal{O}})) \cap h(\mathcal{P}_O) = \varnothing$ since y is closer to $\mathcal{O}$ than any $h(\mathcal{O})$.

In particular, g maps P_g to $P_{g^{-1}}$ and H_g into the closure of $\mathbb{H}^3 \setminus H_{g^{-1}}$. For $x \in P_g$, $d(\mathcal{O}, x) = d(x, g^{-1}(\mathcal{O})) = d(g(x), \mathcal{O})$.

The argument shows that the interior of $\mathcal{P}_{\mathcal{O}}$ cannot contain points of a rotation axis of G; also that there cannot be any overlap in the interiors in the G-orbit of $\mathcal{P}_{\mathcal{O}}$.

(b) If x is an interior point of a face $f' \subset P_g$, we have $d(g(x), \mathcal{O}) = d(x, \mathcal{O}) < d(x, g^{-1}h(\mathcal{O}))$, so long as $h \neq \mathrm{id}$. Thus $g(x) \in P_{g^{-1}}$ also lies in a face.

On the other hand, no conjugate hgh^{-1} can also be a face pairing transformation. Instead, hgh^{-1} is a face pairing transformation of $\mathcal{P}_1 = h(\mathcal{P}_{\mathcal{O}})$.

(c) There cannot be different faces f_1, f_2 with the property that $g_1(f_1) = g_2(f_2) = f$. For $g_1(\mathcal{P}_{\mathcal{O}})$ is exterior to $\mathcal{P}_{\mathcal{O}}$ but adjacent to f. The transformation g_2^{-1} maps f to f_2 and necessarily sends $g_1(\mathcal{P}_{\mathcal{O}})$ back to $\mathcal{P}_{\mathcal{O}}$. Thus $h = g_2^{-1} g_1$ maps $\mathcal{P}_O$ onto itself so $g_1 = g_2$ and hence $f_1 = f_2$.

We conclude that the faces of $\mathcal{P}_{\mathcal{O}}$ are arranged in mutually disjoint (except for perhaps a common edge) isometric pairs.

(d) *The edge relations.* Choose an edge e_1 and then one of the two faces sharing e_1, say f_1. A face pairing transformation g_1 sends the partner face f_1' to $f_1 = g_1(f_1')$ and $g_1(\mathcal{P}_{\mathcal{O}})$ is adjacent to $\mathcal{P}_{\mathcal{O}}$ along f_1. An edge e_2 of f_1' is sent by g_1 to e_1. A special case is when g_1 is elliptic and e_1 is contained in its axis of rotation. Then the partner face f_1' and f_1 both share the edge e_1. If g_1 has order m, the m polyhedra $\mathcal{P}_{\mathcal{O}}, g_1(\mathcal{P}_{\mathcal{O}}), \cdots, {g_1}^{m-1}(\mathcal{P}_{\mathcal{O}})$ form a complete cycle of polyhedra, sharing the edge e_1, each sharing a face with the adjacent polyhedra. In this case the *edge relation* determined by e_1 is simply ${g_1}^m = \text{id}$.

Otherwise there is a face $f_2 \neq f_1$ that shares with f_1' the edge e_2. Its partner face f_2' is sent by some g_2' to $f_2 = g_2(f_2')$. There is an edge e_3 of f_2' that g_2 sends to e_2. Note that the three polyhedra $\mathcal{P}_{\mathcal{O}}, g_1(\mathcal{P}_{\mathcal{O}}), g_1g_2(\mathcal{P}_{\mathcal{O}})$ are arranged in cyclic order about the edge e_1. Successive polyhedra share a face.

Next take the face $f_3 \neq f_2'$ that also shares e_3 and find its partner and the face pairing map $g_3(f_3') = f_3$. To our cyclic arrangement about e_1 we can add one more, $g_1g_2g_3(\mathcal{P}_{\mathcal{O}})$. Keep going. The process will necessarily end after a finite number of steps. We will arrive at f_k with the property that f_k' shares e_1 with f_1. At this point the polyhedra $\mathcal{P}_{\mathcal{O}}, g_1(\mathcal{P}_{\mathcal{O}}), \ldots, g_1g_2 \ldots, g_k(\mathcal{P}_{\mathcal{O}})$ are arranged in cyclic order about e_1. Furthermore the transformation $h = g_1g_2 \ldots, g_k$ fixes the edge e_1. There are two possibilities.

The first is that $g_1 \cdots g_k = \text{id}$, that is, the final polyhedron in the cycle, namely $g_1g_2 \ldots, g_k(\mathcal{P}_{\mathcal{O}})$, coincides with $\mathcal{P}_{\mathcal{O}}$. The *edge relation* determined by e_1 is $h = \text{id}$. The sequence of edges $e_1, \ldots, e_k = e_1$, is called an *edge cycle*. Had we started instead with a different edge e_j in the cycle, its edge relation is conjugate to that for e_1. The dihedral angles corresponding to the edges in the cycle sum to 2π.

The second possibility is that $h = g_1 \cdots g_k$ is an elliptic transformation fixing the edge e_1, and $k \geq 1$ is the smallest number with this property. If h has order m then for $\mathcal{P}^* = \mathcal{P}_{\mathcal{O}} \cup g_1(\mathcal{P}_{\mathcal{O}}) \cup \cdots \cup g_1g_2 \cdots g_k(\mathcal{P}_{\mathcal{O}})$, the collection $\mathcal{P}^*, h(\mathcal{P}^*), \ldots, h^{m-1}(\mathcal{P}^*)$ is a nonoverlapping cyclic ordering of km polyhedra about e_1. The edge relation associated with e_1 is $h^m = \text{id}$. The sequence of edges $e_1, e_2, \ldots, e_k = e_1$ forms an *elliptic edge cycle*. Each edge in the cycle is contained in the rotation axis of an elliptic element conjugate to h. The sum of the dihedral angles about e_1 of the polyhedra $\mathcal{P}_{\mathcal{O}}, g_1(\mathcal{P}_{\mathcal{O}}), \ldots, g_k(\mathcal{P}_{\mathcal{O}})$ must be $2\pi/m$.

By adjoining to $\mathcal{P}_{\mathcal{O}}$ the polyhedra which share a face with $\mathcal{P}_O$, and then those that share just an edge, we can completely surround $\mathcal{P}_{\mathcal{O}}$ by other polyhedra of its orbit. A vertex v of $\mathcal{P}_{\mathcal{O}}$ will be shared exactly by those polyhedra that are part of the edge cycles about the edges of $\mathcal{P}_{\mathcal{O}}$ that end at v.

(e) *The G-orbit of* $\mathcal{P}_{\mathcal{O}}$ *covers* $\mathbb{H}^3$. For suppose to the contrary that the orbit does not cover $y \in \mathbb{H}^3$. Consider the geodesic segment $[\mathcal{O}, y]$. At most a finite number of elements in the orbit can intersect this segment. There is a point $w \in [\mathcal{O}, y]$ such that w lies on the boundary of some element of the orbit, but no point closer to y does. But we know we can completely surround any element $h(\mathcal{P}_{\mathcal{O}})$ of the orbit by other neighbors sharing a face or edge. Therefore $w = y$ and y is covered, after all.

(f) As a consequence we can assert that the rotation axis of each elliptic in G contains a segment which is conjugate to an edge of $\mathcal{P}_{\mathcal{O}}$. For if not, the rotation axis of some conjugate g would meet the interior of $\mathcal{P}_{\mathcal{O}}$. But then g could not send $\mathcal{P}_{\mathcal{O}}$ into its exterior, a contradiction.

The rotation axis of an elliptic g is the line of intersection of the two planes $P_{g^{\pm 1}}$. If the rotation axis of a primitive elliptic contains an edge e of $\mathcal{P}_{\mathcal{O}}$, the two faces sharing e must necessarily be contained in the planes $P_{g^{\pm 1}}$. Therefore g is a face pairing transformation, and no conjugate can also be face pairing.

It is time to bring up a special case: Suppose f' is contained in $P_{g^{-1}}$ for g elliptic of order two. Then $g(P_{g^{-1}}) = P_g = P_{g^{-1}}$, and $g(f') = f'$. The face f' is divided in two parts by the rotation axis of g and application of g interchanges the two parts. To incorporate this special case into our general theory, we must allow any segment of the rotation axis that meets $P_{\mathcal{O}}$ to be counted as an edge of $\mathcal{P}_{\mathcal{O}}$, and regard f' itself as the union of two adjacent faces.

(g) *The presentation of* G. In the G-orbit of $\mathcal{P}_{\mathcal{O}}$, the first generation of polyhedra consists of those that share an edge with $\mathcal{P}_{\mathcal{O}}$. The second generation consists of those which share an edge with a member of the first generation. The n-th generation consists of polyhedra that share an edge with the $(n-1)$ generation but not with a member of an earlier generation. It is clear that any given compact subset of $\mathbb{H}^3$ is covered by the polyhedra in a sufficiently high generation. This shows that the face pairing transformations of $\mathcal{P}_{\mathcal{O}}$ generate G: any $g \in G$ can be written as a composition of face pairing transformations by following a connected union of polyhedra in the orbit, beginning with $\mathcal{P}_{\mathcal{O}}$ and ending with $g(\mathcal{P}_{\mathcal{O}})$.

A small sphere about a vertex v is subdivided into circular polygons by its intersection with the polyhedra sharing v.

Consider the graph Γ_e formed by the union of the edges of the polyhedra in the orbit of $\mathcal{P}_{\mathcal{O}}$. This may or may not be connected. But any simple loop in $\mathbb{H}^3 \setminus \Gamma_e$ is homotopic to a finite product of tiny circles about edges, connected by an arc to the base point of the fundamental group. Furthermore each edge is conjugate to an edge of $\mathcal{P}_{\mathcal{O}}$. This translates into the statement that all relations in G are generated by the edge relations of $\mathcal{P}_{\mathcal{O}}$. For any relation in the generators $g_1 g_2 \cdots g_k = \mathrm{id}$ corresponds to a loop in the complement of Γ_e.

(h) If $\Omega(G) \neq \varnothing$, set $\mathcal{P}_* = \overline{\mathcal{P}} \cap \Omega(G)$. We claim that the G-orbit of $\mathcal{P}_*$ covers $\Omega(G)$ without overlap on the interiors. But this is clear from the fact the orbit of $\mathcal{P}_{\mathcal{O}}$ covers $\mathbb{H}^3$ without overlap. In general $\mathcal{P}_*$ is not connected. The sides of $\mathcal{P}_*$ are outer edges of faces of $\mathcal{P}_{\mathcal{O}}$, and the vertices of $\mathcal{P}_*$ are endpoints of edges. Thus the sides of $\mathcal{P}_*$ are arranged in pairs where the side pairing transformations also generate G.

(i) Finally if $\mathcal{P}_{\mathcal{O}}$ is truncated by intersection with $B_R(\mathcal{O})$, the ball of radius R about $\mathcal{O}$, the truncated faces of $\mathcal{P}_{\mathcal{O}}$ are still arranged in pairs, with the same pairing transformations as before. This is because if a point x in a face σ is distance R from $\mathcal{O}$ and $g : \sigma \to \sigma'$ is the face pairing transformation, then $g(x) \in \sigma'$, being equidistant from $\mathcal{O}$ and $g(\mathcal{O})$, is also distance R from $\mathcal{O}$. $\square$

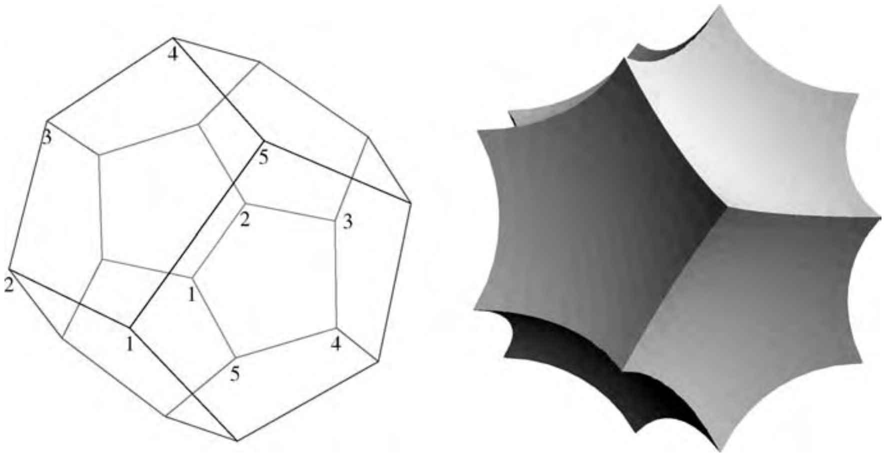


Fig. 3.1. A regular hyperbolic dodecahedron with 72° dihedral angles (right). There is a Möbius transformation that maps each face to the opposite face with a $\frac{3}{10}$ clockwise twist. These generate a kleinian group. The quotient manifold is called the Seifert–Weber dodecahedral space. Its first homology group vanishes. The combinatorial pattern of the identifications is shown on the left.

The Ford fundamental region and polyhedron

In this section we will work with the upper half-space model. For the basic facts about isometric circles and planes we refer back to Section 1.7. They are defined for all elements $\neq$ id in a group provided ∞ is not a limit point. So long as $\Omega(G) \neq \varnothing$, we can replace G by a conjugate if necessary so that $\infty \in \Omega(G)$. Then every element has a well defined isometric circle and isometric plane which is the hemisphere that rises from the isometric circle.

For $g \in G$, let $\mathcal{E}(g)$ and $\mathcal{E}^*(g)$ denote the closure of the exterior of the isometric circle for g and the isometric hemisphere rising from that circle, respectively. In line with our penchant to define "fundamental regions" as relatively closed sets, we define the *Ford region* or *isometric fundamental region* F and the *Ford polyhedron* or *isometric fundamental polyhedron* $\mathcal{F}$ as the following relatively closed sets:

$$F = \left(\bigcap_{g\in G} \mathcal{E}(g)\right) \cap \Omega(G), \quad \mathcal{F} = \bigcap_{g\in G} \mathcal{E}^*(g).$$

The isometric polyhedron is a limiting case of Dirichlet polyhedra. For if g does not fix ∞, as $\mathcal{O} \to \infty \in \partial\mathbb{H}^3$, H_g converges to the complement of the isometric hemisphere for g (see Exercise 3-4). From this we see that $\mathcal{P}_\mathcal{O}$ converges to $\mathcal{F}$ uniformly on compact subsets of $\mathbb{H}^3$.

The polyhedron $\mathcal{F}$ (as well as F) has all the properties listed in Proposition 3.5.1:

Lemma 3.5.2. *If $\infty \in \Omega(G)$, the isometric fundamental polyhedron $\mathcal{P}_\infty(G) = \mathcal{F}$ in the upper half-space model is well defined and is the limit of the Dirichlet polyhedra $\mathcal{P}_\mathcal{O}$ as $\mathcal{O} \to \infty$.*

If instead $\mathcal{O} = \infty$ *is a parabolic fixed point of* G, *the convex polyhedron* $\widetilde{\mathcal{P}}_\infty$ *exterior to the isometric planes of all* $g \in G$, $g(\infty) \neq \infty$, *is periodic with respect to the stabilizer* $\mathrm{Stab}(\infty)$, *while the elements of the* G*-orbit of* $\widetilde{\mathcal{P}}_\infty$ *correspond to the* G*-cosets of* Stab_∞.

The second statement follows from Proposition 1.5.5(7). It is sometimes a very useful object to consider, in spite of the periodicity. This is especially true when there is only one cusp.

The intersection with $\Omega(G)$ of the euclidean closure $\overline{\mathcal{P}}_\infty$ is the isometric region $F = \overline{\mathcal{P}}_\infty \cap \Omega(G)$. It may have isolated points, as a church steeple rising toward $\Omega(G)$ from $\mathcal{P}_\infty$. This subtlety is of concern only if one desires precise information about F itself because a neighborhood of an isolated point is covered by a finite number of elements in the orbit of F. The Ford region itself is not necessarily connected and its intersection with a component of $\Omega(G)$ may not be connected. Certainly it is not connected if $\Omega(G)$ is not connected. In any case the orbit of F tiles the region of discontinuity $\Omega(G)$ without interior overlap.

The interior of F is characterized by the property that for any $g \neq \mathrm{id} \in G$, $|g'(z)| < 1$ for $z \in F$. Therefore among all tiles in the orbit of F, it is F itself that is largest, in view of the formula

$$\iint_{g(F)} du\, dv = \iint_F |g'(z)|^2 dx\, dy < \iint_F dx\, dy = \infty.$$

Since $|g'(z)|^2 = O(|z|^{-2})$ as $|z| \to \infty$, the intermediate integral is automatically finite. The inequality becomes more meaningful if $\Omega(G)$ has a bounded component Ω and we replace F by $F \cap \Omega$ and G by $\mathrm{Stab}(\Omega)$.

We can now prove the following extended form of the universal horodisk theorem.

Corollary 3.5.3. *Suppose* $\Omega \subset \mathbb{C}$ *is invariant under a nonelementary group* G *without elliptics and containing the translation* $T(z) = z + 1$. *Assume that* T *is determined by a puncture in* Ω/G. *Then there exists* $M \geq 0$ *such that* Ω *contains one of the two half-planes* $\{z : |\operatorname{Im} z| > M\}$. *Its image under any* $g \in G$ *that does not fix* ∞ *is disjoint.*

Proof. The isometric circle of any $g \in G$ that does not fix ∞ has its center $g^{-1}(\infty)$ on $\partial\Omega$. As a consequence of the universal horoball property, its isometric circle has radius not exceeding one. Any $g \in G$ that does not fix ∞ sends the exterior of its isometric circle onto the interior of that of g^{-1}.

The assumption on the quotient insures there is a horodisk σ at $+i\infty$ in the hyperbolic metric on Ω and $\partial\sigma$ is an open analytic arc of period one. In the fundamental strip $S = \{z : 0 \leq \operatorname{Re} z < 1\}$, the set $\partial\sigma \cap S$ is uniformly bounded above. Therefore σ contains a half-plane.

Without the assumption on the quotient, the conclusion would be false, as we will well understand when we discuss deformations and pinching: If T is not determined by a puncture in the quotient, it acts in Ω as if it were loxodromic. □

Poincaré's Theorem

A particular consequence of Proposition 3.5.1 is the *local finiteness* of the G-orbits of $\mathcal{P}_\mathcal{O}$ in $\mathbb{H}^3$ and $\overline{\mathcal{P}}_\mathcal{O} \cap \Omega(G)$ in $\Omega(G)$: Any neighborhood of a point intersects only a finite number of elements of the orbit.

It is possible to have a polygon or polyhedron that seems to have the properties of a fundamental region, yet it does not have the local finiteness property. A nice example is presented in [Mumford et al. 2002, Project 7.1] (another example is [Beardon 1983, 9.2.5]): Consider the group generated by the two parabolics $A(z) = z+3$ and $B(z) = 2z/(3z+2)$, which acts in the upper and lower half-plane. The element A maps the circle $C_1 = \{|z+1/2| = 1/2\}$ onto $C_2 = \{|z-1| = 1\}$, sending the inside of C_1 onto the outside of C_2. The element B maps the line $C_3 = \{\operatorname{Re} z = -1\}$ onto $C_4 = \{\operatorname{Re} z = 2\}$ sending the right side of C_3 onto the right side of C_4. The group $G = \langle A, B\rangle$ is discrete and preserves the upper and lower half-planes. In fact G is a variation on the modular group M_2 of Exercise 2-9. The element $A^{-1}B$ is loxodromic with fixed points $-2, -1$. Therefore $\lim_{n\to+\infty}(A^{-1}B)^n(C_3)$ is the circle $\{|z + 3/2| = 1/2\}$. The quotient $\mathbb{H}^2/G$ is conformally equivalent to a twice punctured disk. The region exterior to C_1, C_2 and between C_3 and C_4 has the properties of a fundamental region, except it is not locally finite. It has an edge which ends at a fixed point of a loxodromic element but which is not itself preserved by that element; the projection of the edge to the quotient spirals into the corresponding geodesic without meeting it.

It is also possible to have a polyhedron that seems to be a fundamental polyhedron but the face pairing transformations do not generate a discrete group. Take a convex euclidean quadrilateral Q with no two sides parallel. Find the two affine mappings $A_i(z) = a_i z + b_i$ that map one side to its opposite side and send Q to a polygon $A_i(Q)$ that does not overlap Q except along a side. The two elements generate a nondiscrete group in $\mathbb{C}$. In the upper half-space model, above Q rises a chimney Q^*. The transformations A_i act in $\mathbb{H}^3$ and are hyperbolic isometries pairing opposite faces of Q^*, as required of face pairing transformations. Yet the group they generate is not discrete. What went wrong? This example is from [Epstein and Petronio 1994].

Still, if we start with a convex polyhedron Q^* with the properties (1), (3) of Proposition 3.5.1, the face pairing transformations will in general generate a discrete group for which Q^* is a (locally finite) fundamental region. This is called *Poincaré's Theorem*. One must be particularly careful in understanding the orbit of the ends of the polyhedron on $\partial\mathbb{H}^3$. While this is often self-evident if there are a finite number of faces, in the presence of infinitely many faces special care must be taken. For the definitive analysis, valid in all dimensions, see [Epstein and Petronio 1994].

Note that there are perfectly good fundamental regions that are neither Dirichlet nor isometric fundamental regions. Simple examples are most fundamental parallelograms for discrete, rank two groups of translations. Another example is the modular group M_2 of Exercise 2-6; there one pair of circles are isometric circles, but the other pair are not. However one might think of the fundamental region as a truncated Ford polygon because ∞ is a parabolic fixed point.

The Cayley graph corresponding to a Dirichlet polyhedron

The dual graph Λ associated with $\mathcal{P}_{\mathcal{O}}$ is constructed as follows. Draw a geodesic from $\mathcal{O}$ to each point $g(\mathcal{O})$ where $g(\mathcal{P}_{\mathcal{O}})$ shares a face with $\mathcal{P}_{\mathcal{O}}$. Then draw geodesics to the centers of the polyhedra of the G-orbit that share faces with the first generation, and so on. We get an infinite connected graph Λ, embedded in $\mathbb{H}^3$. If each edge cycle has length 3, which is true in the generic case, then there is a geodesic triangle transverse to each edge. These geodesic triangles are the 2-simplices of the graph.

The graph Λ is equivariant under G. Its projection is therefore a graph $\Lambda_* \subset \mathcal{M}(G)$. The edges of the graph project to simple loops from $\pi(\mathcal{O})$. These loops generate the fundamental group $\pi_1(\mathcal{M}(G); \mathcal{O})$. The 2-cells generate the relations in $\pi_1(\mathcal{M}(G); \mathcal{O})$.

For the general definition of Cayley graphs see Exercise 2-17. If $\mathcal{P}_{\mathcal{O}}$ has an infinite number of faces, Λ is perhaps not useful. In contrast, the abstract Cayley graph for G does not suffer under the same handicap.

Additional remarks

Wielenberg [1981] has given examples showing that a polyhedron may be the fundamental polyhedron for more than one group; different pairings of faces give rise to different groups. An example of this phenomenon for fuchsian groups is in Exercise 2-13.

R. Riley over many years developed a computer program to test whether a group given by generating matrices is discrete [Riley 1983]. In effect, it tests for discreteness using Jørgensen's inequality and the universal horoball property, and then it tries to construct an isometric fundamental polyhedron. If successful, the program can read off the presentation of the group.

Jørgensen [1973] has completely analyzed the isometric fundamental polyhedron for cyclic loxodromic groups $\langle T \rangle$ in terms of the trace parameter, using the normalization of Exercise 1-34. The polyhedron can have an arbitrarily large number of faces; large numbers of faces arise when the trace with $|\mathrm{tr}(T)| < 2$ tangentially approaches 2. (When $|\mathrm{tr}(T)| > 2$, the isometric circles of $T^{\pm 1}$ are disjoint.) The combinatorial arrangement of faces is completely described in terms of $\mathrm{tr}(T)$. Moreover, either F is the region bounded by the isometric circles of $T^{\pm 1}$ (when $|\mathrm{tr}(T)| \geq 2$) or it is the closure of a simply connected domain with either four or six sides. Wada [$\geq$ 2007a] wrote a computer program showing the structure of the isometric fundamental polyhedron as a function of the trace.

Jørgensen also analyzed the Ford fundamental polyhedra of once-punctured torus groups G in terms of the combinatorics of the faces. He shows how, starting with the side pairing transformations of the Ford regions (typically bounded by six circular arcs) on the two components of $\Omega(G)$, the sequence of face pairing transformations of the Ford polyhedron can be read off. This study has been important to this day because this class is the simplest nontrivial class of groups, depending on only two complex parameters. Besides important applications in its own right, especially to two-bridge knots [Akiyoshi et al. 1999], it serves as a test bed for more general situations. For details of Jørgensen's analysis see [Jørgensen 2003; Akiyoshi et al. 2003], [Akiyoshi et al. 2005].

3.6 Geometric finiteness

The importance of the class of geometrically finite groups lies in the fact that the class corresponds to the manifolds $\mathcal{M}(G)$ which are "essentially" compact [Marden 1974a]. The longstanding conjecture that geometrically finite groups are dense in all finitely generated kleinian groups has recently been proved. (This will be discussed in some detail in Sections 5.4–5.6.) The precise definition and characterization is as follows ("pairs of punctures" and "solid pairing tubes" will be explained below):

Assume G is nonelementary and, unless stated otherwise, has no elliptics.

Theorem 3.6.1 [Marden 1974a; 1977]. *Given a base point $\mathcal{O} \in \mathbb{H}^3$, $\mathcal{P}_{\mathcal{O}}(G)$ has a finite number of faces if and only if the quotient manifold $\mathcal{M}(G)$ is compact except perhaps for a finite number of rank one and rank two cusps, and the rank one cusps correspond to pairs of punctures on $\partial\mathcal{M}(G)$.*

If the condition holds for one base point $\mathcal{O}$, it holds for any choice of base point.

Corollary 3.6.2. *A manifold $\mathcal{M}(G)$ is geometrically finite if and only if (i) the punctures on $\partial\mathcal{M}(G)$ are arranged in pairs such that each pair determines a solid pairing tube, and (ii) the result of removing the interiors of all solid pairing tubes and solid cusp tori is a compact manifold $\mathcal{M}_0(G)$.*

A group that has a finite sided Dirichlet polyhedron is called *geometrically finite*. Correspondingly, a *geometrically finite manifold* is one that is the quotient of such a group. The term also applies to orbifolds.

Schottky groups and finitely generated fuchsian groups (Section 2.7) are examples of geometrically finite groups with $\Omega(G) \neq \varnothing$. Alternate characterizations of geometric finiteness are given in terms of the convex core in Section 3.10.3, the conical limit points in Exercise 3-18, and the Hausdorff dimension of the limit set in Exercise 3-20.

The term was coined by Leon Greenberg. After Ahlfors' announcement of his finiteness theorem, the next thought was that a Dirichlet region in $\mathbb{H}^3$ for a finitely generated kleinian group had to have a finite number of faces, as is the analogous case in $\mathbb{H}^2$ for fuchsian groups. This hope was decisively dashed when Greenberg pointed out that this is not the case for the "degenerate" groups (Chapter 5) discovered by Bers on the boundary of Teichmüller space [Greenberg 1966; Marden 1974a]. This was the first indication that $\mathbb{H}^3$ really matters.

In contrast, consider the following interesting fact, a consequence of the Ahlfors Finiteness Theorem. For proofs see [Beardon and Jørgensen 1975], [Greenberg 1977], and Exercise 3-31 below.

The boundary $\overline{\mathcal{P}}_{\mathcal{O}} \cap \Omega(G)$ "at ∞" of a Dirichlet polyhedron $\overline{\mathcal{P}}_{\mathcal{O}}(G)$, or the Ford fundamental region $F(G)$, where G is finitely generated, has a finite number of sides.

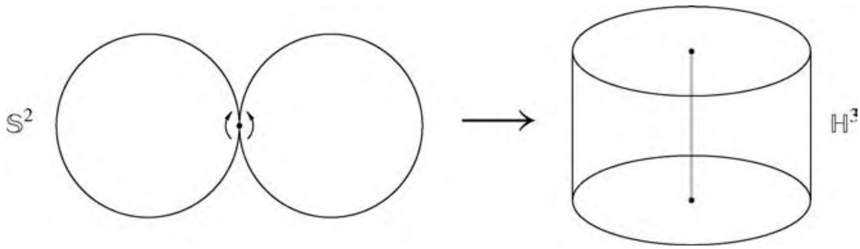


Fig. 3.2. Solid pairing tube for a rank one cusp.

The special role of parabolics. Let $\zeta \in \partial\mathbb{H}^3$ be a parabolic fixed point of G and Stab_ζ the parabolic subgroup fixing ζ. We have called ζ a *rank one* or *rank two cusp* if Stab_ζ has one or two generators respectively. Associated with ζ is its universal horoball $\mathcal{H}$, whose "size" depends only on a least length generator (Section 3.2). For $T \notin \mathrm{Stab}_\zeta$, $T(\mathcal{H}) \cap \mathcal{H} = \varnothing$ while $T(\mathcal{H}) = \mathcal{H}$ for $T \in \mathrm{Stab}_\zeta$.

The geometric structure associated with a rank two cusp is the same for all hyperbolic manifolds, even those with nonfinitely generated fundamental groups. Embedded in $\mathcal{M}$ is a one-parameter family of solid cusp tori (Section 3.2) for every conjugacy class of rank two parabolic subgroups. The universal horoball property assures us that if we choose the solid cusp tori to come from horoballs properly contained in the universal horoballs, those corresponding to different conjugacy classes have mutually disjoint closures.

If the interiors of the solid cusp tori interiors are removed from $\mathcal{M}(G)$, there results a manifold with the same fundamental group but with a number of torus boundary components. These are in addition to the components of $\partial\mathcal{M}(G)$, none of which can be tori. Every noncyclic abelian subgroup of $\pi_1(\mathcal{M}) \cong G$ arises by an injection into $\pi_1(\mathcal{M})$ of the fundamental group of a cusp torus. A particular horoball associated with a rank two cusp can be chosen to be of maximal size in that its boundary torus is just tangent to itself; this is not necessarily true of the universal horoball. The set of volumes of these maximal solid cusp tori is an invariant of the particular hyperbolic structure.

If there are elliptics sharing the fixed point ζ then instead of the solid cusp torus there will be an object homeomorphic to $S' \times [0, \infty)$ where S' is a sphere with three or four cone points.

Rank one cusps in a geometrically finite group are associated with a very particular geometric structure that may not appear in a general group. The geometric structure is much stronger than the mere existence of a horosphere and solid cusp tube (Section 3.2); the solid cusp tube must be directly related to two punctures on the boundary of the manifold. In a geometrically finite group, there corresponds a *pair of punctures* p_1, p_2 on $\partial\mathcal{M}(G)$, uniquely associated with the conjugacy class of the cusp: If c_1, c_2 are small circles in $\partial\mathcal{M}(G)$ retractable to p_1, p_2, there is a *pairing cylinder* C in $\mathcal{M}(G)$, which is a cylinder, closed in $\mathcal{M}(G)$, and bounded by c_1 and c_2. It bounds a

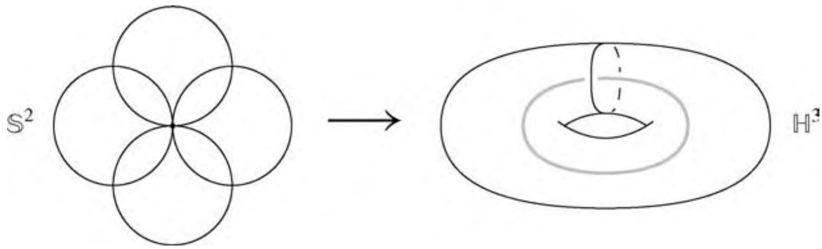


Fig. 3.3. Solid cusp torus for a rank two cusp.

subregion of $\mathcal{M}(G)$, called a *solid pairing tube*, which is homeomorphic to $C \times (0, 1]$ (and retractable to a cusp). The solid pairing tubes corresponding to the different conjugacy classes of rank one cusps can be chosen to be mutually disjoint in the geometrically finite manifold $\mathcal{M}(G)$.

Let T be a parabolic generator of an element of the conjugacy class that represents the cusp. The circles c_1, c_2 can be chosen so that the pair lifts to round circles in $\Omega(G)$ mutually tangent at the fixed point ζ of T; see Corollary 3.5.3. Such a pair of circles is called a *double horocycle* at ζ, even though this is an abuse of terminology if the components of $\Omega(G)$ containing them are not round disks on $\mathbb{S}^2$.

Suppose the fixed point ζ is shared by an order two elliptic. then instead of the solid pairing cylinder there will be an object of the form $D^* \times [0, \infty)$. The subset $D^* \times \{0\}$ of $\mathrm{Int}(\mathcal{M}(G))$ is a disk with one puncture or cone point.

Consider a fuchsian manifold $\mathcal{M}(G)$ with G acting on the upper and lower half-planes; every puncture on one component of $\partial\mathcal{M}$ is paired with a puncture on the other. Suppose $T : z \mapsto z+1$ is a least length generator of a rank one parabolic subgroup. For $b > 1$, $\{z \in \mathbb{C} : \mathrm{Im}\, z = \pm b\}$ is a pair of horocycles at the fixed point ∞. These project to "circles" about a pair of punctures. Let $P_\pm \subset \mathbb{H}^3$ denote the vertical planes rising from them and consider the vertical slab $Q = \{(z, t) \in \mathbb{H}^3 : -b \leq \mathrm{Im}\, z \leq b,\ t > 0\}$ they bound. Truncate Q by the half-space $K = \{(z, t) : t \geq a > 1\}$. The relative boundary in $\mathbb{H}^3$ of the resulting tunnel $Q \setminus K$ projects to a pairing tube. This explicit construction suggests how solid pairing tubes can be created in general — there does not seem to be a canonical construction.

Proof of Theorem 3.6.1. We continue to assume that G has no elliptics — elliptics will be dealt with at the end of the proof. Assume that $\mathcal{P}_\mathcal{O}$ has a finite number of faces. The set $F = \overline{\mathcal{P}}_\mathcal{O} \cap \mathbb{S}^2$ consists of a finite number of finite sided circular polygons and perhaps a finite number of isolated points. By a vertex of F we mean a point which is on the edge of at least two different faces of $\mathcal{P}_\mathcal{O}$; either the circles bounding two faces cross at v, or they are tangent at v.

We claim that if a point $x \in F$ is contained in the closure of only finitely many elements $\mathcal{P}_1 = \mathcal{P}_\mathcal{O}, \mathcal{P}_2, \ldots, \mathcal{P}_k$ of the orbit $G(\mathcal{P}_\mathcal{O})$, then $x \in \Omega(G)$. To verify this statement (as in [Greenberg 1966]), construct a horosphere $\mathcal{H}_x$ at x so small in

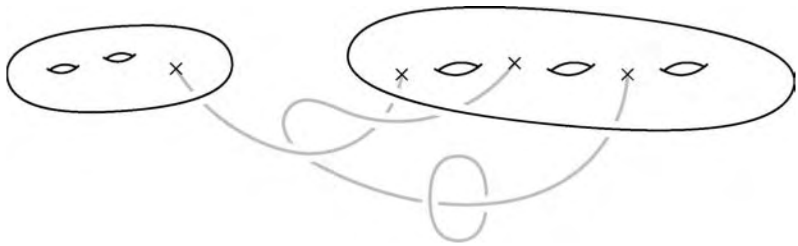


Fig. 3.4. Schematics of a geometrically finite manifold. Each gray line joining a pair of $\times$'s indicates a pairing tube; the gray closed curve indicates a rank two cusp.

spherical diameter that it intersects only the polyhedra $\{\mathcal{P}_j\}$. This means that $\mathcal{H}_x$ is partitioned into sectors, each of which lies in some $\mathcal{P}_j$. A neighborhood of x on $\mathbb{S}^2$ is likewise partitioned. Therefore $x \notin \Lambda(G)$.

If there are no parabolics, each vertex of F is completely surrounded by a finite number of elements of the orbit of F; no vertex cycle can result in a parabolic while a loxodromic cannot fix a point on closure of $\mathcal{P}_{\mathcal{O}}$. Likewise the edges in $\mathbb{H}^3$ are also completely surrounded. Therefore $\mathcal{M}(G)$ is compact.

Now consider an ideal point $x \in \overline{F} \cap \Lambda(G)$. Because there are only a finite number of faces, x lies in the boundary of infinitely many elements $\{\mathcal{P}_j\}$ of the G-orbit $G(\mathcal{P}_{\mathcal{O}})$. Of the infinitely many faces of the $\{\mathcal{P}_j\}$ that contain x on their boundary, infinitely many are images of the same face of $\mathcal{P}_{\mathcal{O}}$ by elements of $\mathrm{Stab}_x \subset G$. All these transformations must be parabolic. For if $T \in \mathrm{Stab}_x$ were loxodromic, and if P were a plane with x in its boundary, than the limit points of P under the cyclic group $\langle T \rangle$ is the axis of T. This is impossible for a plane containing a face in the G orbit of $\mathcal{P}_{\mathcal{O}}$. So x is the common fixed point of a rank one or two parabolic subgroup and the Dirichlet region $\mathcal{P}_x = \mathcal{P}_{\mathcal{O}}(\mathrm{Stab}_x)$ contains $\mathcal{P}_{\mathcal{O}}$. If Stab_x has rank one, $\mathcal{P}_x$ is the region bounded by two hyperbolic planes which are tangent at x. If Stab_x has rank two, $\mathcal{P}_x$ is a chimney of four or six faces rising to x.

We have to consider in more detail the case where Stab_x is rank one. In $\mathbb{S}^2$, choose two circles tangent at x that bound a strip S_x whose Stab_x-orbit is all $\mathbb{C}$. For example, if $x = \infty$ and Stab_x is generated by $z \mapsto z + 1$, we can choose the strip $S_x = \{0 \leq \mathrm{Re}\, z \leq 1\}$. The intersection with S_x of a small neighborhood of x must lie in F since boundaries of faces of $\mathcal{P}_0$ cannot accumulate to x within S_x. This shows that there is a double horocycle at x with respect to Stab_x. In other words with respect to $\partial\mathcal{M}(G)$, x supports a pair of punctures.

We conclude that $\mathcal{M}(G)$ is compact except for a finite number of solid cusp tori and solid cusp tubes with respect to pairs of punctures.

Conversely, if $\mathcal{M}(G)$ has the "essential compactness" just described, we claim that $\mathcal{P}_{\mathcal{O}}$ has a finite number of faces. Otherwise, where in $\mathcal{M}(G)$ would the projection of an infinite number of faces $\{\pi(f_j)\}$ accumulate? We know there can be no accumulation point within $\mathcal{M}(G)$.

Suppose infinitely many $\pi(f_j)$ were in the interior C of a solid cusp torus. A face $\pi(f_i) \subset C$ does not separate C. Therefore there is a simple loop in C, not retractable to a point, joining one side to the other. This loop determines an element of the fundamental group $\pi_1(C)$, which is a rank two abelian group. Because not more than one pair of faces can be paired by elements of a cyclic subgroup, the projection of at most two faces can lie inside C, a contradiction. The same argument applies to the interior of a solid pairing tube. We conclude that $\mathcal{P}_O$ has a finite number of faces.

Corollary 3.6.2 follows from our argument.

We will indicate how the corresponding theorem for orbifolds can be derived from the theorem for manifolds. By Selberg's Lemma (page 68), there is a torsion free normal subgroup H of finite index. Let $G = \bigcup_{i=1}^{N} g_i H = \bigcup_{i=1}^{N} H g_i$ be a decomposition by distinct cosets. Then $\mathcal{P}^* = \bigcup_{i=1}^{N} g_i(\mathcal{P}_{\mathcal{O}}(G))$ serves as a fundamental domain for H. Although it may not be connected, it has the properties of $\mathcal{P}_{\mathcal{O}}$, in particular the faces are arranged in pairs with respect to H. For example, if $(f, g(f))$ is a pair of faces of $\mathcal{P}_{\mathcal{O}}(G)$ then the $2N$ faces $\{g_i(f), g_i g(f)\}$ are arranged in N pairs under H. Now $g_i g = h g_j$ for some j and $h \in H$ — because $G = Gg = \bigcup g_i H g = \bigcup g_i g H = \bigcup H g_i$. Therefore the faces $g_i(f)$ and $g_j(f)$ are paired by $h \in H$. Also we know that $h_1 g_j \neq h_2 g_k$ for $k \neq j$, $h_1, h_2 \in H$. In effect, $\mathcal{P}_{\mathcal{O}}(H)$ is made up of N copies of $\mathcal{P}_{\mathcal{O}}(G)$. We conclude that G is geometrically finite if and only if H is as well. (The picture at orbifold cusps is more complicated if there are elliptics that share the parabolic fixed points.) □

Lemma 3.6.3 [Thurston 1986b]. *If G is geometrically finite and $\Omega(G)$ is nonempty, every finitely generated subgroup is also geometrically finite.*

A proof is indicated in Exercise 3-7. Without the assumption that $\Omega(G) \neq \varnothing$, the statement would be false in general, as we will later see in Section 6.1.

Finite volume

Lemma 3.6.4 [Wielenberg 1977]. *If* $\mathrm{Vol}(\mathcal{M}(G)) < \infty$, *then G is geometrically finite.*

Proof. Again we may assume that G has no elliptics. Consider the ε-thick part $\mathcal{M}(G)^{\text{thick}}$ (with ε chosen as in Proposition 3.4.1). The surface area of a cusp cylinder coming from a rank one cusp is infinite. Therefore a small neighborhood in the thick part would have infinite volume. So G cannot have any rank one cusps. On the other hand the volume of each ε-solid cusp torus is not less than $2\varepsilon^2|\tau|\sin\theta \geq \sqrt{3}\varepsilon^2$ by Exercise 2-10, so there are at most a finite number of them. If the thick part were not compact there would be an infinite sequence $x_n \in \mathcal{M}(G)^{\text{thick}}$ which are centers of mutually disjoint ε balls. Therefore the volume of $\mathcal{M}(G)$ would have to be infinite, which is not the case. □

3.7 Three-manifold surgery

In this section we will present what is needed from 3-manifold topology for direct application to hyperbolic manifolds. For a rigorous treatment of the aspects of topology that we are using, we refer to [Hempel 1976] or [Jaco 1980].

Dehn's Lemma and the Loop Theorem. *Let S be a boundary component of an orientable 3-manifold M^3. Suppose $\gamma \subset S$ is a simple loop homotopic to a point within M^3 but not within S. Then γ is the boundary of an essential disk.*

Suppose a nonsimple loop $\gamma \subset S$ is homotopic to a point in M^3 but not in S. Given a neighborhood $N_\gamma \subset S$ of γ, there a simple loop $\gamma_0 \subset N_\gamma$ that bounds an essential disk $D \in \mathcal{M}(G)$.

An *essential disk* is an embedded closed disk $D \subset M^3$ such that $D \cap \partial M^3 = \partial D$, where ∂D is not homotopic to a point in ∂M^3. We call a loop $\gamma \subset S$ *nontrivial* if it is not homotopic to a point within S. When obtaining a disk from application of Dehn's Lemma and the Loop Theorem, we will automatically choose one that is essential. A boundary component that supports an essential disk is called *compressible*.

The equivariant version is also useful:

Equivariant Dehn's Lemma and the Loop Theorem [Meeks and Yau 1981]. *Suppose X is a finite group of automorphisms of some $\mathcal{M}(H)$ with compressible boundary. Then there is a set of mutually disjoint compressing disks whose members are permuted by G and project injectively to $\mathcal{M}(H)/X$.*

In our applications M^3 is a smooth, oriented manifold, and γ can also be chosen to be smooth. There is an important generalization:

Cylinder Theorem. *Suppose $\gamma_1, \gamma_2 \subset \partial M^3$ are disjoint nontrivial simple loops that are freely homotopic in M^3 but not within ∂M^3. There is an essential cylinder embedded in M^3 bounded by γ_1 and γ_2.*

Suppose instead that the freely homotopic loops are not simple but $\gamma_i \subset N_i \subset \partial M^3$, where the neighborhoods N_1 and N_2 are disjoint. There are simple loops $\gamma_i' \subset N_i$ that bound an essential cylinder in M^3.

That two loops are *freely homotopic* means that there is a continuous mapping of an annulus A into M^3 sending the boundary components of A to the two loops. Another way of describing free homotopy is as follows: γ_1 is freely homotopic to γ_2 if and only if there is an arc α from any given point $O_1 \in \gamma_1$ to any given point $O_2 \in \gamma_2$ such that γ_1 is homotopic to $\alpha^{-1}\gamma_2\alpha$ (here we are composing curves from right to left).

Two disjoint simple loops that are freely homotopic in ∂M^3, but neither is homotopic to a point in ∂M^3, bound a (topological) annulus in ∂M^3.

An *essential cylinder* is a closed cylinder $C \subset M^3$ such that $C \cap \partial M^3 = \partial C$, the boundary components of C are not homotopic to points in M^3, and C cannot be homotoped (relative to $\partial\mathcal{M}(G)$, that it is allowed to slide along $\partial\mathcal{M}^3$) to an annulus in $\partial\mathcal{M}^3$. When obtaining a cylinder from application of the Cylinder Theorem, we will automatically choose one that is essential.

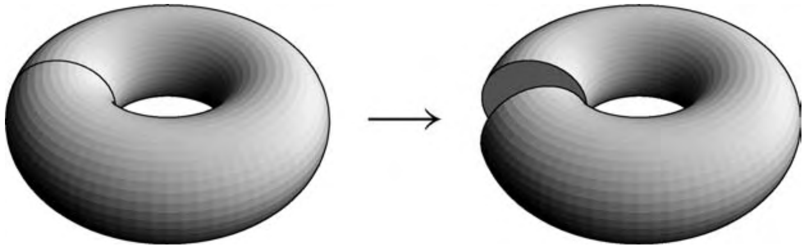


Fig. 3.5. Cutting a solid torus along a compressing disk results in a topological ball.

In the case of a kleinian manifold $\mathcal{M}(G)$ we will add the following requirement to the definition: For C to be called an essential cylinder, *it cannot bound a solid pairing tube*. Here we are regarding a pairing cylinder as homotopic into the boundary.

It is possible that a simple loop $\gamma \in \partial\mathcal{M}(G)$ may be a boundary component of two or more homotopically distinct essential cylinders which are disjoint, except for sharing the common boundary γ. On the other hand,

A simple nontrivial loop on a cusp cylinder or cusp torus cannot be freely homotopic to a loop either on a cusp cylinder or cusp torus corresponding to a different cusp.

Application of Dehn's Lemma and the Loop Theorem

If a component Ω of $\Omega(G)$ is not simply connected, there is a simple loop $\gamma^* \in \Omega$ which separates its boundary components. Of course γ^* is homotopic to a point if we move it into $\mathbb{H}^3$. Its projection $\gamma \subset R = \Omega/\operatorname{Stab}(\Omega)$ is a closed loop, perhaps not a simple loop, which is not homotopic to a point in R, but is homotopic to a point in $\mathcal{M}(G)$. Dehn's Lemma and the Loop Theorem say that there is a simple loop $\gamma' \in R$ which bounds an essential disk in $\mathcal{M}(G)$.

A component R of $\partial\mathcal{M}(G)$ is *incompressible* if the inclusion $\pi_1(R) \hookrightarrow \pi_1(\mathcal{M}(G))$ is injective. Our argument shows that R is incompressible if and only if all the components of $\Omega(G)$ which lie over R are simply connected. Otherwise R is called *compressible*. If all the boundary components are incompressible, the manifold $\mathcal{M}(G)$ is called *boundary incompressible*.

More generally, an orientable surface S embedded in $\mathcal{M}(G)$ is called *incompressible* if it is not a topological disk and if the inclusion $\pi_1(S) \hookrightarrow \pi_1(\mathcal{M}(G))$ is injective. This means that every loop in S which is homotopic to a point in $\mathcal{M}(G)$ is already homotopic to a point in S. Otherwise there is a simple loop in S bounding an essential disk whose interior lies in $\mathcal{M}(G)\setminus S$ [Jaco 1980, III.8]. The surface S is incompressible if and only if each lift over it in $\mathbb{H}^3$ is simply connected.

An essential disk D in $\mathcal{M}(G)$, $\partial D \subset \partial\mathcal{M}(G)$, is called a *compressing disk*. It either divides $\mathcal{M}(G)$ into pieces M_1, M_2 or $M_1 = \mathcal{M}(G) \setminus D$ is connected. In the first case the fundamental group of $\mathcal{M}(G)$ splits into a free product: $\pi_1(\mathcal{M}) = \pi_1(M_1) * \pi_1(M_2)$ and correspondingly G splits: $G = G_1 * G_2$ (van Kampen's Theorem). In the second case let γ be a simple loop from an origin $O \in M_1$ that crosses D once. Then

$\pi_1(\mathcal{M}(G)) = \langle \pi_1(M_1), \gamma \rangle$ or $G = \langle G_1, T \rangle$ where $TG_1T^{-1} = G_1$, $T \notin G_1$ (G is an HNN-extension of G_1).

The subgroup $\pi_1(M_i)$ corresponds to a conjugacy class of subgroups of G — take a lift M_i^* of M_i to $\mathbb{H}^3$ and let G_i denote its stabilizer. There are one or more copies of the compressing disk D in the relative boundary of M_i^* in $\mathbb{H}^3$. These lifted disks bound a topological half-space of $\mathbb{H}^3$ not containing M_i^*. Adding these half-spaces to M_i^* gives back all $\mathbb{H}^3$. Moreover the half-spaces project injectively into $\mathcal{M}(G_i)$.

Now starting with some manifold $\mathcal{M}(G)$, the process of repeated insertion of compressing disks, which we can take to be mutually disjoint, terminates after a finite number of steps. We end up with a union of manifolds that are either balls or are boundary incompressible (see [Hempel 1976] or [Jaco 1980]). For example, if we start with a handlebody of genus g coming from a Schottky group, after cutting it along g mutually disjoint disks none of which divide the handlebody we will end up with a topological ball. See Exercise 3-11.

Here is a way of reversing the process of cutting $\mathcal{M}(G)$ by an essential disk: Choose disjoint, closed, round disks D_1, D_2 in $\Omega(G)$. Choose them small enough that each projects injectively into $\partial\mathcal{M}(G)$ and they remain disjoint there. Let σ_1, σ_2 denote the hyperbolic planes rising from the circles $\partial D_1, \partial D_2$. Let σ_i^- denote the half-space adjacent to D_i and σ_i^+ the other half-space. Choose any Möbius transformation T which has the property that $T(\sigma_1) = \sigma_2$ and $T(\sigma_1^+) = \sigma_2^-$. We see that $G^* = \langle G, T \rangle$ is a discrete group: T conjugates all the action of G in σ_1^+ to the action of TGT^{-1} in σ_2^-. Of course the operation is duplicated over the full orbit $G(D_i)$. The associated manifold $\mathcal{M}(G^*)$ is obtained from $\mathcal{M}(G)$ as follows. Down in $\mathcal{M}(G)$ we have the disks $D_i \subset \partial\mathcal{M}(G)$, and the planes σ_i which lie in the interior of $\mathcal{M}(G)$ except for their boundaries and bound balls (here we are using the same notation for the projections). Let M denote the result of removing from $\mathcal{M}(G)$ the two half-spaces σ_i^-. The action by T forms a new hyperbolic manifold $\mathcal{M}(G^*)$ from M by gluing σ_1 to σ_2. In $\mathcal{M}(G^*)$, $\sigma_1 \equiv \sigma_2$ is an essential disk which does not separate.

The procedure works equally well if we have two manifolds $\mathcal{M}(G_i)$ and take a disk in each boundary. In this case the new essential disk will divide the manifold. This process we have described is an example of *Klein–Maskit combination theory*, developed by Klein and refined and extended by Maskit [1988]; see [Marden 1974a] for the manifold interpretation. See also Exercise 3-8.

For the following result, see for example [Waldhausen 1968].

Proposition 3.7.1. *Suppose M^3 is a compact, orientable and irreducible 3-manifold. If $\pi_1(M^3) = A * B$, $A, B \neq \mathrm{id}$, is a free product of subgroups, there exists a compressing disk bounded by a simple loop in ∂M^3.*

Remark 3.7.2. In calculating the genus of the boundary of a 3-manifold in terms of its fundamental group the following simple fact is very useful. Suppose $\alpha, \beta \in \partial M^3$ are two 1-cycles with nonzero intersection number. Then at most one of them can be homologous to zero, or, in particular, homotopic to a point in M^3. Thus if the fundamental group has N generators so that its first homology group has at most

N generators, the total genus of the boundary is at most N. In particular, if G has no parabolics, $\partial\mathcal{M}(G)$ has at most $N/2$ components. Similar arguments give useful estimates for the topology of the boundary [Marden 1971; 1974a]. For example, if G is a g-generator free group ($g \geq 2$) and $\mathcal{M}(G)$ is compact with $\partial\mathcal{M}(G)$ a closed surface of genus g, then $\mathcal{M}(G)$ is a handlebody.

Equivariant extensions $\partial\mathcal{M} \to \mathcal{M}$

Often we will be in the position of having a group G and a quasiconformal deformation $F : \Omega(G) \to \Omega(H)$ that induces an isomorphism $\varphi : G \to H$. Such a map is called equivariant; it is the lift of a quasiconformal map $f : \partial\mathcal{M}(G) \to \partial\mathcal{M}(H)$ which (i) sends puncture pairs to puncture pairs, and (ii) sends compression loops to compression loops.

We will spell out in terms of given basepoints how the boundary map f respects the isomorphism $\varphi : \pi_1(\mathcal{M}(G);\, O) \to \pi_1(\mathcal{M}(H);\, O')$. On each boundary component R of $\partial\mathcal{M}(G)$, choose a basepoint p, and then choose the basepoint $f(p) \in f(R)$. To each loop $\alpha \subset R$ with basepoint p corresponds a loop $f(\alpha) \subset f(R)$ with basepoint $f(p)$. Upon joining the loops to the basepoints O, O' by auxiliary arcs, we get inclusion homomorphisms $\pi_1(R;\, p) \hookrightarrow \pi_1(\mathcal{M}(G);\, O)$ and $\pi_1(f(R);\, f(p)) \hookrightarrow \pi_1(\mathcal{M}(H);\, O')$ with kernels $K = \pi_c(R)$, $K' = \pi_c(f(R))$. There are a finite number of mutually disjoint simple compression loops on R such that the kernel K is the least normal subgroup $\pi_c(R) \subset \pi_1(R)$ generated by these (see the Maskit Planarity Theorem, p. 76). In turn the map f induces an isomorphism between the images of the inclusions.

We want to find a quasiconformal extension to $f : \mathcal{M}(G) \to \mathcal{M}(H)$. Although no "canonical" method seems available, the extension can be done by topological means (extension is not always possible in the geometrically infinite case).

Suppose first that $\mathcal{M}(G)$ is compact. According to [Hempel 1976, Theorem 13.9 and Corollary 13.7], f is homotopic on $\partial\mathcal{M}(G)$ to a homeomorphism f_1 which has an extension to a homeomorphism between the manifolds $f_1 : \mathcal{M}(G) \to \mathcal{M}(H)$. In turn f_1 is homotopic to a diffeomorphism $f_2 : \mathcal{M}(G) \to \mathcal{M}(H)$, [Munkres 1960]; f_2 is automatically quasiconformal. We can choose a lift F_2 of the new f_2 to $\mathbb{H}^3 \cup \Omega(G)$ so that its restriction to $\Omega(G)$ induces φ and is homotopic to F. But now, applying [Gehring 1962], F_2 has an φ-equivariant quasiconformal extension to all of $\mathbb{S}^2$.

If there are parabolics we have to replace the manifolds by the compact manifolds resulting from the removal of the solid pairing tubes and the solid cusp tori and extend the extension back to the original manifolds.

For applications it suffices to replace (F, f) by (F_2, f_2). However it is nicer to apply the stronger result Theorem 3.7.3 below.

It is shown in Exercise 3-34 that the original F itself has a homeomorphic extension to $\mathbb{S}^2$ satisfying $F(\zeta) = F_2(\zeta)$ for all $\zeta \in \Lambda(G)$. In fact, the extension of F is quasiconformal on all $\mathbb{S}^2$ by Theorem 3.14.6. This puts us in a position to apply Theorem 3.7.4(iii) below. We end up with a most satisfying result as follows:

Theorem 3.7.3. *Assume that G is geometrically finite and F is a quasiconformal mapping $\Omega(G)$ onto $\Omega(H)$ that induces an isomorphism $\varphi : G \to H$. Then F is the restriction of an equivariant quasiconformal map of $\mathbb{S}^2$ which extends to an equivariant quasiconformal mapping $F : \mathbb{H}^3 \cup \mathbb{S}^2 \to \mathbb{H}^3 \cup \mathbb{S}^2$. The mapping F then projects to a quasiconformal mapping $f : \mathcal{M}(G) \to \mathcal{M}(H)$.*

Now suppose $\mathcal{M}(G)$ is not necessarily geometrically finite. We start afresh with a quasiconformal mapping $F : \mathbb{S}^2 \to \mathbb{S}^2$ that induces an isomorphism $\varphi : G \to H$ satisfying $F(g(z)) = \varphi(g)F(z)$ for all $g \in G$, $z \in \mathbb{S}^2$. If the restriction of F is conformal $\Omega(G) \to \Omega(H)$, or if $\Omega(G) = \varnothing$, F is is Möbius and the two groups are conjugate. Here we are applying Theorem 5.6.6 or, if $\Omega(G) = \varnothing$, Corollary 3.13.4.

One general approach is the following. It is based on a canonical method of Douady and Earle to extend a quasiconformal automorphism F of the $(n-1)$-sphere (when $n = 2$ such a map is called *quasisymmetric*) to a surjective mapping of n-ball for $n \geq 2$. The extension is equivariant if F is so. That is, if F satisfies $F \circ g(z) = \varphi(g) \circ F(z)$ for all $z \in \mathbb{S}^{n-1}$ and $g \in G$ for *any* Möbius group G and an isomorphism to another group $\varphi : G \to H$, then its extension is also equivariant with the same φ. On the other hand the extension is guaranteed to be a homeomorphism only when $n = 2$, or when the complex dilatation of the boundary mapping is sufficiently small (see [McMullen 1996, p. 231]). Fortunately in the case $n = 3$ a modification suggested by Pekka Tukia (personal communication) allows one to get a homeomorphism of the ball without any restrictions. This modification is based on the fact [Ahlfors 1966, p. 100] that in dimension 2, given $\varepsilon > 0$, a quasiconformal mapping can be factored into the composition $F = F_n \circ F_{n-1} \circ \cdots \circ F_1$ of a finite number of equivariant quasiconformal mappings each of whose complex dilatations satisfies $\|\mu_k\|_\infty < \varepsilon$. This is done by taking $\mu_k = (k/n)\mu$ for sufficiently large n, where μ is the complex dilatation of F. In consistent normalizations, denote the solution of the corresponding Beltrami equation by g_k. Then set $F_k = g_k \circ g_{k-1}^{-1}$. The Douady-Earle extension is then applied to each factor F_k resulting in an extension to a equivariant homeomorphism of $\mathbb{H}^3$.

The weakness of this approach is that the extension is not known to be quasiconformal or even bilipschitz. There is an alternate approach by integrating an extension of a vector field on $\mathbb{S}^2$. This method is suggested in in [Thurston 1979, Chapter 11; Reimann 1985] and carried out in [McMullen 1996, Corollary B.23]. The possibilities are itemized below.

Theorem 3.7.4 (Basic Extension Theorems). *Suppose G, H are arbitrary kleinian groups and $F : \mathbb{S}^2 \to \mathbb{S}^2$ is a K-quasiconformal mapping that induces an isomorphism $\varphi : G \to H$. Then*:

(*i*) [Douady and Earle 1986; Tukia 2005] *The map F has an equivariant extension to $\mathbb{H}^3$ that is a homeomorphism which also induces φ; its projection $f : \mathcal{M}(G) \to \mathcal{M}(H)$ is an orientation preserving homeomorphism.*

(*ii*) [Tukia 1985c] *The map F has an equivariant (L, a)-quasiisometric extension for some $L = L(K)$, $a = a(K)$; its projection $f : \mathcal{M}(G) \to \mathcal{M}(H)$ is a (L, a)-quasiisometric mapping.*

(*iii*) [McMullen 1996, Corollary B.23] *The map F has an equivariant extension to a $K^{3/2}$-bilipschitz diffeomorphism of $\mathbb{H}^3$; its projection $f : \mathcal{M}(G) \to \mathcal{M}(H)$ is a $K^{3/2}$-bilipschitz diffeomorphism (and hence quasiconformal).*

A mapping f of $\mathbb{H}^3$ is (L, a)-quasiisometric if there exist finite constants $1 \leq L$ and $a \geq 0$ such that in the hyperbolic metric

$$\frac{1}{L}d(x, y) - a \leq d(f(x), f(y)) \leq Ld(x, y) + a.$$

Thus a quasiisometric map need not be continuous but at long range it is essentially bilipschitz. Like quasiconformal maps of $\mathbb{H}^3$ [Gehring 1962], quasiisometric maps can be extended to $\partial\mathbb{H}^3 \equiv \mathbb{S}^2$ and the extension is a quasiconformal map of $\mathbb{S}^2$. If in addition it is a homeomorphism, it will automatically be quasiconformal (but quasiconformal maps are not automatically bilipschitz). See Exercise 3-19.

3.8 Quasifuchsian groups

A *quasifuchsian group* G is the quasiconformal deformation (page 80) of a fuchsian group Γ. The purpose of this section is to characterize this class of groups by the topology of $\mathcal{M}(G)$.

Assume first we have a finitely generated kleinian group G with $\Omega(G) = \Omega_1 \cup \Omega_2$ such that G preserves Ω_1, Ω_2. We will show that G is quasifuchsian. By the Ahlfors Finiteness Theorem, the quotients $\Omega_i/G = R_i$ are closed surfaces with at most a finite number of punctures and branch points.

Each component must be simply connected. Otherwise there would a simple loop α in Ω_1, say, that separates its boundary. This would force Ω_2 to make a choice of which component of $\mathbb{S}^2 \setminus \alpha$ to lie in. Whichever it chose, its boundary could not be the full limit set, a contradiction.

Choose a fuchsian group Γ and quasiconformal mappings $f_1 : \mathrm{UHP}/\Gamma \to R_1$, $f_2 : \mathrm{LHP} \to R_2$ that lift to $F_1 : \mathrm{UHP} \to \Omega_1$, $F_2 : \mathrm{LHP} \to \Omega_2$. We also need the reflection $J : z \mapsto \bar{z}$. We must choose these maps so that the orientation reversing map $H = F_2 \circ J \circ F_1{}^{-1} : \Omega_1 \to \Omega_2$ satisfies $H \circ g = g \circ H$ for all $g \in G$. Once f_1 is chosen, the homotopy type of f_2 is determined by this requirement. In particular there is an isomorphism $\phi : \Gamma \to G$ for which $F_i \circ \gamma = \phi(\gamma) \circ F_i$, $i = 1, 2$, for all $\gamma \in \Gamma$. The complex dilation μ of F_1, F_2 is a Beltrami differential on UHP $\cup$ LHP. Define μ to vanish on $\mathbb{R} \cup \{\infty\}$. Solve the Beltrami equation. There results a quasiconformal map F of $\mathbb{S}^2$ that conjugates Γ to a quasifuchsian group G'. Because they solve the same Beltrami equation, $F \circ F_i^{-1}$ is conformal on Ω_i, $i = 1, 2$ and induces an isomorphism $\phi' : G \to G'$. Anticipating Theorem 3.13.3, ϕ' is a conjugation. So after renormalizing F, we may assume F restricts to F_1 and F_2.

We will now show that if $\Omega(G)$ has two invariant components Ω_1, Ω_2 then $\Omega(G)$ has only the two components Ω_1, Ω_2 as claimed in Lemma 2.4.2(iii).

From the perspective of $\partial\mathcal{M}(G) = R_1 \cup R_2$ there is an "identity" isomorphism $j : \pi_1(R_1) \to \pi_1(R_2)$ that comes from identification of the action of $g \in G$ on Ω_1 with its action on Ω_2: Fix basepoints $O_i^* \in \Omega_i$ and a geodesic τ^* in $\mathbb{H}^3$ that connects them. Given $g \in G$, a simple arc $\gamma_i^* \subset \Omega_i$ from O_i^* to $g(O_i^*)$ projects to a loop $\gamma_i \in R_i$ from the point $O_i = \pi(O_i^*)$, $i = 1, 2$. Then $\gamma_1^* \sim g(\tau^*)\gamma_2^*\tau^{*-1}$ and the projections to $\mathcal{M}(G)$ satisfy $\gamma_1 \sim \tau\gamma_2\tau^{-1}$; γ_1 is freely homotopic to γ_2 in $\mathcal{M}(G)$. This gives the isomorphism j. A change of basepoints will give the same isomorphism j if the connecting arc τ^* is correspondingly adjusted.

Applying the Cylinder Theorem, given a simple loop $\gamma_1 \subset R_1$ there is a simple loop $\gamma_2 \subset R_2$ that bounds an essential cylinder within $\mathcal{M}(G)$.

A very similar situation arises for punctures and cone points. If $g \in G$ is parabolic, since each Ω_i is simply connected, its fixed point corresponds to a puncture in each of R_i. According to Corollary 3.5.3, its fixed point supports a horocycle in both Ω_1 and Ω_2. From this, we can construct a solid pairing tube in $\mathcal{M}(G)$ pairing the two punctures. Also each elliptic transformation has one fixed point in each component and its axis of rotation extends from one to the other, analogous to the situation for a parabolic.

The argument proceeds as follows. Suppose first that there are no elliptics or parabolics. Consider a simple closed geodesic c in the hyperbolic metric on R_1. There is a corresponding geodesic c' in R_2 such that c, c' are the boundary components of a cylinder in $\mathcal{M}$. If d is a simple geodesic in R_1 crossing c exactly once and $d' \subset R_2$ corresponds to d, the two cylinders can be adjusted so that they are transverse to each other within $\mathcal{M}(G)$—they intersect in a single arc.

Now take a chain of $2g$ simple geodesics $\{c_i\}$ in R_1, where g is the genus, such that c_i crosses c_{i-1} and c_{i+1} while c_{2g} crosses c_{2g-1} and c_1, but otherwise the geodesics are mutually disjoint. The complement of their union in R_1 is simply connected. Insert cylinders so that within $\mathcal{M}(G)$ each is transverse to its neighbors but disjoint from the others. Let M denote the complement in $\mathcal{M}(G)$ of the union of the cylinders. The interior of M can only be a ball because it is bounded by a topological 2-sphere. This establishes the product structure for this case.

In the general case, choose mutually disjoint solid cusp pairing tubes for the pairs of punctures and solid tubes about the rotation axes. Connect the union of the geodesics $\{c_i\}$ with the circles about the punctures and branch points in R_1 so the result bounds a simply connected region. Then extend this to connect within $\mathcal{M}$ the union of the cylinders with the cylinders about punctures and branch points to once again get a complementary region M bounded by a topological 2-sphere. The argument is completed as before. So there is an (orientation preserving) homeomorphism $\mathcal{M}(G) \to R_1 \times [0, 1]$.

Appealing in addition to Theorem 3.7.3, we conclude that there is a quasiconformal mapping $F : \mathbb{S}^2 \to \mathbb{S}^2$, taking $\mathrm{UHP} \cup \mathrm{LHP}$ to $\Omega_1 \cup \Omega_2$, such that:

(i) there is an isomorphism $\varphi : \Gamma \to G$ such that $(F \circ \gamma)(z) = (\varphi(\gamma) \circ F)(z)$ for all $z \in \mathbb{S}^2$ and $\gamma \in \Gamma$.
(ii) $F \circ J \circ F^{-1}$ is an involution of G.
(iii) F projects and extends to a quasiconformal map $f : \mathcal{M}(\Gamma) \to \mathcal{M}(G)$.

Simultaneous uniformization

Suppose Γ is a fuchsian group acting again in the upper and lower half-planes and then in upper half-space. The orientation reversing involution $J_0 : (z, t) \in \mathbb{H}^3 \mapsto (\bar{z}, t)$ interchanges the upper and lower half-planes, UHP and LHP, and pointwise fixes the vertical plane P rising from $\mathbb{R}$. It satisfies $J_0 \circ \gamma = \gamma \circ J_0$ for all $\gamma \in \Gamma$. The projection J_{0*} to $\mathcal{M}(\Gamma)$ is an anticonformal mapping that exchanges the two boundary components, and pointwise fixes P/Γ.

A quasifuchsian deformation G of Γ is induced by a quasiconformal map (see Section 3.6.3) $f : \mathbb{H}^3 \cup \partial\mathbb{H}^3 \to \mathbb{H}^3 \cup \partial\mathbb{H}^3$ that satisfies $f \circ \gamma(x) = \theta(\gamma) \circ f(x)$ for all $x \in \mathbb{H}^3 \cup \partial\mathbb{H}^3$, for an isomorphism $\theta : \Gamma \to G$. The map f projects to a homeomorphism $f_* : \mathcal{M}(\Gamma) \to \mathcal{M}(G)$. Hence $J_* = f_* \circ J_{0*} {f_*}^{-1}$ is an orientation reversing involution of $\mathcal{M}(G)$, exchanging its two boundary components and inducing the identity automorphism of $\pi_1(G)$.

It is customary to refer to the boundary component UHP$/\Gamma$ as the *top* boundary component of $\mathcal{M}(\Gamma)$ and LHP$/\Gamma$ as the *bottom* and correspondingly for any quasifuchsian deformation of Γ. The following generative result is due to Bers.

Simultaneous uniformization. *Suppose R_{bot}, R^{top} are two Riemann surfaces of finite hyperbolic area and $J : R_{\text{bot}} \leftrightarrow R^{\text{top}}$ is an orientation reversing involution. There exists a quasifuchsian group G, uniquely determined up to Möbius equivalence, such that the top boundary component of $\mathcal{M}(G)$ is conformally equivalent to R^{top}, the bottom conformally equivalent to R_{bot}, and such that J is homotopic to the restriction of J_* to $\partial\mathcal{M}(G)$.*

3.9 Geodesic and measured geodesic laminations

In this section we will introduce the notions of geodesic and measured geodesic laminations in $\mathbb{H}^2$ which are needed to understand the internal structure of hyperbolic manifolds. General references are [Fathi et al. 1979], [Canary et al. 1987], [Bonahon 2001].

It is helpful to think in terms of the disk model of $\mathbb{H}^2$. Let G be a fuchsian group such that $\mathbb{H}^2/G = R$ is a surface of finite hyperbolic area (no elliptics).

Draw any simple closed curve γ from a basepoint $O \in R$ which is not homotopic to a point or a puncture. Choose $O^* \in \mathbb{H}^2$ over O. A lift γ_0^* of γ beginning at O^* terminates at $g(O^*)$ for some $g \neq \text{id} \in G$. The orbit γ^* of γ_0^* under the cyclic group $\langle g \rangle$ is a simple arc in $\mathbb{H}^2$ with end points at the fixed points of g. Since γ^* projects to a simple loop, it will have the property that its orbit under the full group G consists of mutually disjoint arcs.

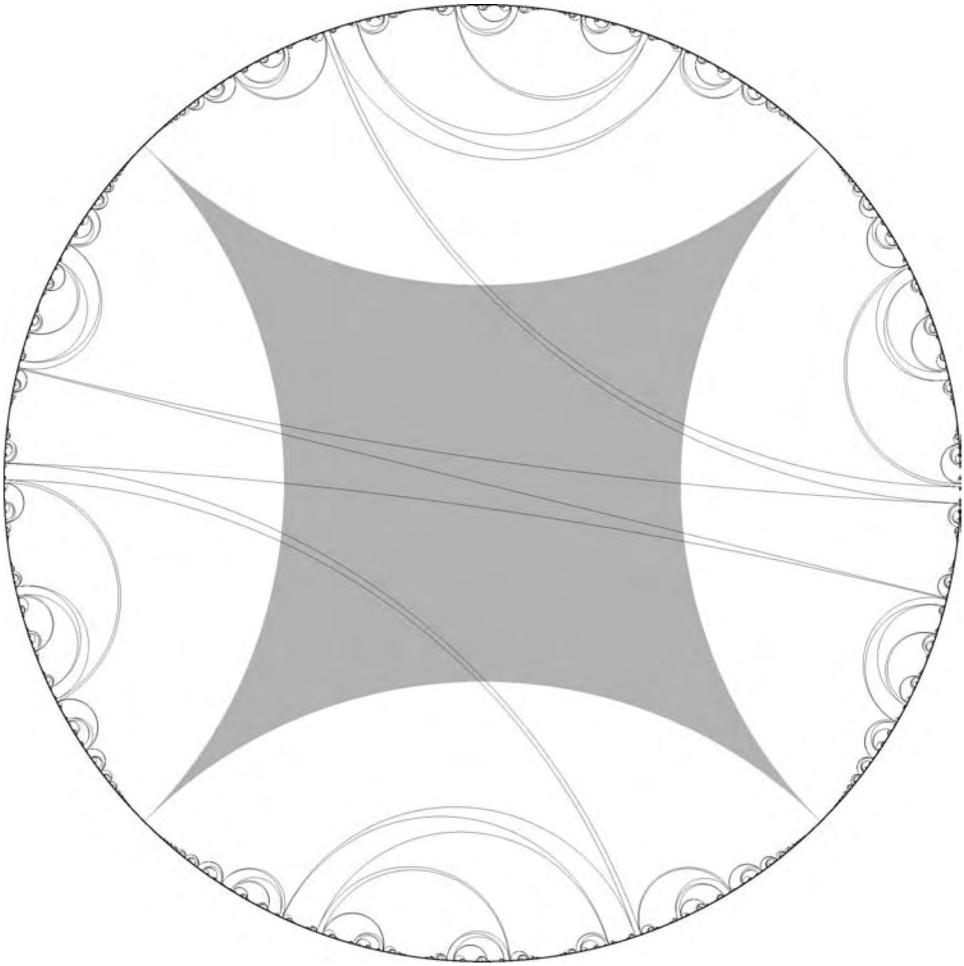


Fig. 3.6. A discrete geodesic lamination consisting of the lifts of a long simple geodesic on a once punctured torus. A fundamental polygon is shaded.

Now consider the axis α^* of g, namely the hyperbolic line between the endpoints of γ^*; α^* projects to a closed loop α on R which is necessarily a simple loop and a geodesic. Furthermore, γ is freely homotopic to α. In Exercise 3-3 we will find that α^* does not penetrate the universal horoballs at parabolic fixed points.

Fix a fundamental polygon P for G (for an explicit example, see Exercise 2-13). Consider a sequence of closed geodesics $\{\alpha_n\}$ that are getting longer and longer, say in terms of a fixed set of generators for $\pi_1(R; O)$. Choose a point $p_n \in \alpha_n$ and a lift so that $p_n^* \in \alpha_n^*$ lies in P; the corresponding axes α_n^* all intersect P. What happens as $n \to \infty$? Since all the axes intersect P (but do not enter the universal horoballs at the cusps of P) we can find a subsequence that converges to a geodesic σ^* in $\mathbb{H}^2$. Necessarily neither endpoint of σ^* is a parabolic fixed point. What about the projection σ to R?

First of all σ, can have no self-intersections, since the orbit of σ^* under G consists of mutually disjoint geodesics. Second, it cannot be a closed geodesic, for $\lim g_n$ cannot exist as a proper Möbius transformation. Therefore σ is a simple geodesic of infinite length on R. As such it has limit points in R — that is there are sequences of points $\{p_n^*\} \in \sigma^*$ which converge to an endpoint so that the projections $\{p_n = \pi(p_n^*)\} \in R$ converge to a point $p \in R$. Such a geodesic σ is called *recurrent*; it keeps returning to a compact set in R.

However we cannot say that in R, $\{\alpha_n\}$ "converges" to σ. For if a different point $p_n' \in \alpha_n$, increasingly far away along α_n from p_n, was lifted to P, the sequence of lifts will not necessarily converge to σ^*, or even to a leaf in the G-orbit of σ^*.

A *geodesic lamination* $\Lambda^* \subset \mathbb{H}^2$ is a *closed* set of mutually disjoint geodesics. Two leaves are allowed to have a common endpoint on $\partial\mathbb{H}^2$. Each component of Λ^* is called a *leaf*. The components of $\mathbb{H}^2 \setminus \Lambda^*$ are called *gaps*. The gaps are ideal polygons, possibly infinite sided, possibly bounded by arcs of $\partial\mathbb{H}^2$.

The space $\mathcal{GL}$ of all geodesic laminations on $\mathbb{H}^2$ is given the topology of Hausdorff convergence: A sequence converges $\Lambda_n^* \to \Lambda^*$ if every neighborhood of Λ^* contains all but a finite number of Λ_n^*, and if $U \in \mathbb{H}^2$ is an open set containing all but a finite number of Λ_n^* then $\Lambda^* \subset U$. With this topology $\mathcal{GL}$ becomes a compact Hausdorff space. Simple closed geodesics are dense in the subspace $\mathcal{GL}_0(R)$ of those geodesic laminations without leaves ending at a puncture (recall that simple closed geodesics cannot penetrate the universal horodisks); see [Canary et al. 1987].

In fact, there is a natural topology on the space of geodesics in $\mathbb{H}^2$ so that it becomes a Möbius band (Exercises 1-3 and 4-15). An individual geodesic becomes a point in the Möbius band while a geodesic lamination becomes a closed pointset.

Assume now that Λ^* is G-invariant. The projection Λ to $R = \mathbb{H}^2/G$ is a closed set of mutually disjoint simple geodesics in R which cover a set of zero area. A leaf $\ell \subset \Lambda$ is *isolated* if every point $z \in \ell$ has a neighborhood whose intersection with Λ consists of a segment of ℓ through z. For example, suppose α, β are disjoint simple closed geodesics on R. There is an isolated geodesic ℓ one end of which spirals infinitely often around one side of α and the other end infinitely often around β. A lift of γ in $\mathbb{H}^2$ will connect one fixed point of a loxodromic over α to a fixed point of a loxodromic over β. Yet ℓ is not a lamination since it is not closed in the space of geodesics. The lamination is $\gamma \cup \alpha \cup \beta$.

A well known result of Birman and Series [1985] is that the set of all simple (but not necessarily closed) geodesics on a finite surface R form a set of Hausdorff dimension one — see Exercise 3-20. An interesting consequence is that almost every geodesic arc $[a, b] \subset R$ is *generic* with respect to simple geodesics in the sense that it is transverse to *every* simple closed geodesic on R [Bonahon 2001, p. 19].

Another consequence is that if Λ has no isolated leaves, then Λ has uncountably many leaves: For any transverse segment τ, $\tau \cap \Lambda$ is totally disconnected; $\tau \cap \Lambda$ is a Cantor set [Bonahon 2001, Prop. 7].

(Mirzakhani [2004] recently established the precise growth of the number $S_X(L)$ of simple *closed* geodesics of length $\leq L$ on a hyperbolic surface X of genus g and

n punctures:

$$S_X(L) \sim nL^{6g+2n-6}, \quad \text{as } L \to \infty,$$

where $n = n(X)$ is a constant depending on X.)

Any lamination can be augmented by additional leaves if necessary so that the gaps are ideal triangles. The projection of gap in $\mathbb{H}^2$ near an ideal vertex either ends at a puncture, or spirals around R without intersecting itself or other gap projections. Since each ideal triangle has area π, there are exactly $2(2g+n-2)$ different gaps (Exercise 3-1).

A given Λ can be covered with open sets $\{U_i\}$ with continuous maps $\phi_i : U_i \cap \Lambda \to X_i \times (0, 1) \subset \mathbb{R}^2$ taking leaves to vertical line segments indexed by $X_i \subset \mathbb{R}$ and so that $\phi_j \circ \phi_i^{-1}(x, y) = (f(x), g(x, y))$ preserves verticality for overlapping neighborhoods.

A lamination Λ is called *minimal* if it has no closed sublaminations. Each geodesic lamination can be decomposed into (i) the union of finitely many infinite isolated leaves whose ends spiral to a minimal sublamination or end at a cusp, and (ii) the union of finitely many minimal sublaminations with the property that every half-leaf is dense in the sublamination. A closed geodesic is in the second category, and a geodesic whose ends are at punctures is in the first.

Measured laminations

The geodesic lamination $\Lambda \subset R$ (and hence its lift $\Lambda^* \subset \mathbb{H}^2$) is called a *measured lamination* if there is a Borel measure μ with support contained in (usually we will assume equal to) Λ. More precisely, each transverse segment τ, with endpoints in gaps, has finite, positive measure $\mu(\tau)$ where the measure depends only on the equivalence class of τ—$\tau_1 \equiv \tau$ if the endpoints of τ_1 are in the same gaps as the endpoints of τ.

We will always require that the measure be *uniformly bounded* in the sense that there is a constant C such that $\mu(\tau) < C$ for all transversals τ of unit length. For such a measure to exist with support in Λ on a punctured surface, no leaves of Λ can end at punctures. For up in $\mathbb{H}^2$, if there is one leaf ending at the fixed point ζ of a parabolic $g \in G$, all the leaves in its $\langle g \rangle$-orbit also end at ζ and are zero asymptotic distance apart. Thus a transverse segment of unit length ultimately crosses infinitely many leaves, forcing any transverse measure to be infinite. Likewise no leaf of Λ can spiral in to a closed geodesic. This would occur if up in $\mathbb{H}^2$, a leaf shares an endpoint with the axis of a loxodromic representing a simple closed geodesic. For this reason, the only isolated leaves of a measured lamination are simple closed geodesics. The minimal gaps are ideal triangles, ideal bigons containing one puncture, and ideal monogons containing one puncture*. An *ideal bigon* is the union of two ideal triangles and an *ideal monogon* one, after it is slit from the puncture to the ideal point. Each ideal triangle has area π, each ideal bigon has area 2π and each monogon has area π. An

* A bigon B (monogon M) on S is a annular region about a puncture whose boundary in S consists of two (one) infinite geodesics whose ends are asymptotic to each other "at ∞". Each lift of B or M is an infinite sided ideal polygon invariant under a cyclic parabolic subgroup.

n sided ideal polygon made up of $n-2$ ideal triangles, or n triangles if it contains one puncture, can also serve as a gap.

Every geodesic lamination Λ has a transverse measure whose support consists of all the minimal sublaminations of Λ [Bonahon 2001, Prop. 9].

The set $\mathcal{ML}(R)$ of *uniformly (locally) bounded, measured laminations* on a finite area surface R of genus g and b punctures is topologized as follows: $(\Lambda_n, \mu_n) \to (\Lambda, \mu)$ if and only if (i) $\Lambda_n \to \Lambda$ in the Hausdorff topology, and (ii) $\lim \mu_n(\tau) = \mu(\tau)$ for all transversals of Λ. However we have to allow the possibility that the support of μ is a proper sublamination of the Hausdorff limit Λ. We also have to allow the zero-lamination with no leaves and zero measure.

If $[\alpha], [\beta]$ are two free homotopy classes of simple loops, their *geometric intersection number* $\iota([\alpha], [\beta])$ is defined to be the minimum number of crossings $\iota(\alpha, \beta)$ of simple loops α, β in their respective free homotopy classes. This minimum is achieved by the geodesics α_g, β_g in the classes. For this reason we will usually use the geodesics to calculate intersection numbers. We set $\iota(\alpha, \alpha) = 0$ so that $\iota(\alpha, \beta) = 0$ implies that either the geodesics α, β are disjoint, or they coincide. Equally we can define the geometric intersection number of two collections of mutually disjoint simple geodesics. And also the intersection number of a geodesic arc with endpoints in gaps and a finite lamination. (see Exercise 2-5 for the torus case).

The most general transverse measure on a finite system of mutually disjoint closed geodesics is obtained by assigning an atomic measure $\alpha(\ell)$ to each leaf. Then for the measured lamination μ we define by linearity

$$\mu(\tau) = \iota(\tau, \mu) = \sum_{\ell \in \Lambda} \alpha(\ell)\iota(\tau, \ell). \tag{3.3}$$

Thurston proved that the measured laminations with support on a simple closed geodesic are dense in $\mathcal{ML}(R)$ [Thurston 1988; Fathi et al. 1979]. In other words, given $(\Lambda, \mu) \in \mathcal{ML}(R)$ there exists a sequence of simple closed geodesics $\{\alpha_n\}$ and a corresponding sequence $\{a_n\}$ of strictly positive numbers such that $\alpha_n \to \Lambda$ and, for any simple loop or arc σ transverse to Λ,

$$\mu(\sigma) = \lim_{n\to\infty} \frac{\iota(\sigma, \alpha_n)}{a_n} = \int_\sigma d\mu. \tag{3.4}$$

Conversely, given a sequence $\{\alpha_m\}$ of simple closed geodesics on a closed surface (or contained in a compact part of a punctured surface), there exists a subsequence $\{\alpha_m\}$ with associated positive numbers $\{a_n\}$ that converges to a nonzero element $(\Lambda, \mu) \in \mathcal{ML}(R)$. Typically, good choices are $a_n = \mathrm{Len}(\alpha_n)$ (Exercise 3-35) or $a_n = \iota(\sigma_0, \alpha_n)$ if σ_0 is a fixed geodesic transverse to all α_n. Here $1/a_n$ is the atomic measure assigned to α_n.

The sequence $\{a_n\}$ is uniquely determined asymptotically, up to a positive multiplicative constant. Namely if $\{a'_n\}$ is another sequence, there exists a constant $C \neq 0$ such that $\lim a_n/a'_n = C$. To see why, set $C(\sigma) = \lim \iota(\sigma, \alpha_n)/a_n$ and $C'(\sigma) =$

$\lim \iota(\sigma, \alpha_n)/a'_n$. For some σ, $C(\sigma) \neq 0$ and for some σ', $C'(\sigma') \neq 0$. Therefore

$$\lim a_n/a'_n = C'(\sigma')/C(\sigma') = C'(\sigma)/C(\sigma) \neq 0, \infty.$$

We emphasize that the support of μ may not be the whole Hausdorff limit Λ To illustrate what can happen, take the lamination consisting of two disjoint simple geodesics α_1, α_2 with assigned integer multiplicities m_1, m_2. We can construct an sequence of simple geodesics $\{\alpha_n\}$ that go nm_1 times around α_1, and nm_2 times around α_2. The Hausdorff limit $\Lambda = \lim \alpha_n$ is a union $\alpha_1 \cup \alpha_2 \cup \ell_1 \cup \ell_2$ where ℓ_1, ℓ_2 are infinite length geodesics each spiraling around one side of α_1 and of α_2. Up in $\mathbb{H}^2$, each end point of a lift of ℓ_i is a fixed point of a transformation determined by one of the closed leaves. If σ is a simple loop transverse to α_1 but not α_2, $\iota(\sigma, \alpha_n)/n \to m_1$ and similarly the limit is m_2 if it is transverse to α_2 but not α_1. If it is transverse to a geodesic cutting $\ell_1 \cup \ell_2$ but not $\alpha_1 \cup \alpha_2$, the limit is zero.

The bottom line is that every geodesic lamination Λ has a transverse measure whose support consists of all the minimal sublaminations of Λ [Bonahon 2001; Otal 1996]. Two minimal laminations with the same (nonzero) transverse measure are identical. In the above example the minimal laminations are α_1 and α_2. The spiraling geodesic cannot be in the support of μ.

The sequence of geodesic lengths $\{\text{Len}(\alpha_n)\}$ also has a limit if it is scaled by the same $\{a_n\}$ as determines μ, namely

$$\text{Len}_\mu(\Lambda) = \lim_{n\to\infty} \frac{\text{Len}(\alpha_n)}{a_n}, \tag{3.5}$$

which exists and is $\neq 0$. It is called the *length* of the measured lamination (Λ, μ). Unlike intersection numbers, the value of L_μ depends on the particular hyperbolic surface R where the measurement is made; it is known to change continuously as the underlying surface is deformed [Kerckhoff 1985]. An intrinsic expression for the length is

$$\text{Len}_\mu(\Lambda) = \iint_R d\ell \times d\mu, \tag{3.6}$$

where $d\ell$ is the hyperbolic length along the leaves of Λ and $d\mu$ is the transverse measure. Thus if Λ is a single simple closed geodesic, and μ is the atomic measure of unit weight, $\text{Len}_\mu(\Lambda)$ is just the geodesic length on R. The integral is obtained from the local product structure of the lamination determined by an open cover, using a partition of unity.

Using a (quasiconformal) homeomorphism f from one finite area surface R to another S, the space $\mathcal{ML}(R)$ can be transferred to $\mathcal{ML}(S)$. Namely let f^* be a lift of f to $\mathbb{H}^2 \equiv \mathbb{D}$. The homeomorphism extends to a homeomorphism also denoted f^* on $\partial\mathbb{D}$. By taking images of endpoints, f^* induces a map of the (lifted) measured lamination over one surface to the other. Then project back to R and S. For this reason, if you have seen $\mathcal{ML}$ on one surface, you have seen it on all surfaces in the deformation space. However metric properties will differ, one surface to the other.

Another way of looking at the length is to take a finite number of transverse arcs $\{\tau_i\}$ that cut Λ into (generally uncountably many) segments of finite length: $\Lambda \setminus \Lambda \cap (\cup \tau_i)$. Associate each arc of Λ with one of the segments τ_i that contains an endpoint. Then integrate the lengths of the finite arcs of Λ with respect to $d\mu$ along the τ_i, From this point of view the continuity of the length function on $\mathcal{ML}(R)$ in terms of change of hyperbolic metric on R follows [Bonahon 2001, p. 21]. Thurston showed that the arcs can be chosen so that the map from the set of simple closed geodesics to $\mathbb{R}^n$ given by $\mu \mapsto (\mu(\tau_1), \dots, \mu(\tau_n)$, induces a homeomorphism to a piecewise linear submanifold of $\mathbb{R}^n$ of real dimension $6g+2b-6$). The manifold is constructed from the set of simple closed geodesics, as Bonahon remarked, by a process akin to the passage from a lattice $\mathbb{Z}^2$ in $\mathbb{R}^2$ to a torus.

Start with $\iota(\sigma, \mu)$ defined for an atomic measure on a simple closed geodesic and a transverse geodesic τ. The intersection number can be extended by continuity to any measured lamination $(\Lambda, \mu) \in \mathcal{ML}$. It can be extended again by continuity to $\iota(\mu, \nu)$ for any pair of measured laminations [Rees 1981]. Specifically, if we write (Λ, μ), (Λ', ν) as limits of simple closed geodesics $\mu = \lim \alpha_n/a_n$, $\nu = \lim \beta_n/b_n$, then

$$\iota(\mu, \nu) = \lim \frac{\iota(\alpha_n, \beta_n)}{a_n b_n};$$

see [Bonahon 1986]. If $\iota(\mu, \nu) = 0$, any component of the support of μ is either identical to a component of the support of ν or disjoint from all of its components.

This generalization of the geometric intersection number remains a topological entity, independent of any particular complex structure the underlying surface R may have.

With an eye on the fact that the sequences $\{a_n\}$ are asymptotically uniquely determined only up to positive constants, it is usually better to use instead the space of *projective measured laminations*

$$\mathcal{PML}(R) = (\mathcal{ML}(R) \setminus 0)/\text{multiplication by scalars}.$$

For then we do not distinguish between measures that are positive multiples of each other. The space $\mathcal{PML}(R)$ is homeomorphic to the sphere $\mathbb{S}^{6g-7}$ (or $\mathbb{S}^{6g+2b-7}$ if there are also b punctures).

There is a theory of measured foliations which is the topological version of measured laminations. Roughly, a measured foliation is a (necessarily singular) foliation with a measure of distances between leaves. Measured foliations are modeled by quadratic differentials; see Exercise 5-24. Every measured foliation comes from a measured lamination by lifting to $\mathbb{H}^2$, showing (noncritical) leaves have endpoints on $\partial\mathbb{H}^2$, and replace the leaves by geodesics with the same endpoints. The converse is also true. The quantitative results depend on taking a pants decomposition of R and classifying the intersection of a measured foliation with each pants (for a taste, see Exercise 3-35). One way of getting lots of nontrivial examples of measured foliations is by means of interval exchange maps, see Exercise 3-36.

Formal introduction to this beautiful and essential subject can be found, for example, in [Thurston 1988; Fathi et al. 1979; Canary et al. 1987; Casson and Bleiler 1988; Bonahon 2001; Otal 1996, Appendix; Marden and Strebel 1984; Matsuzaki and Taniguchi 1998].

Remarks 3.9.1. (i) Each infinite length leaf $\ell \subset R$ of a lamination (with compact support) is *recurrent*: there is a sequence of points $\{\zeta_n\} \subset \ell$ such that along ℓ, $\zeta_n \to \infty$ yet there exists $\zeta \in \ell$ such that in a neighborhood of ζ *in* R, $\lim \zeta_n = \zeta$. Up in $\mathbb{H}^2$, this says that given a lift ℓ^*, there is a sequence of (mutually disjoint) lifts ℓ_n^* which converge to ℓ^* as euclidean circular arcs.

(ii) Two measured laminations in $\mathbb{H}^2$ whose set of leaves have the same combinatorics and the same transverse measures are usually not Möbius equivalent. For example, they may be lifts of finite laminations on two different surfaces where the distances between leaves differ (compare with Theorem 3.11.3).

(iii) We have seen how a lamination consisting of two or more mutually disjoint simple geodesics has many projectively inequivalent transverse measures. Yet there are geodesic laminations which support *only one* projective class of measures [Masur 1982]; such measured laminations are called *uniquely ergodic*. Uniquely ergodic laminations Λ have the minimality property that Λ is not the union of proper sublaminations. Uniquely ergodic laminations are dense in all measured laminations. Yet it is a subtle business to determine if a particular lamination is uniquely ergodic. The pair of laminations fixed by a pseudo-Anosov automorphism of a surface (see Exercise 5-6) does have this property [Thurston 1988]. The analogous result on a square torus is a famous theorem of Hopf, which says that the projection to the quotient torus of a line of irrational slope in the square lattice in $\mathbb{C}$ is equally distributed on the torus.

A measured lamination (Λ, μ) on a finite area surface S is called *arational* if each complementary component of Λ is an ideal polygon, possibly containing a single puncture [Otal 1996]. Consequently there are at most $4g+2n-4$ gaps for an arational lamination, where $n \geq 0$ is the number of punctures. An arational lamination is cut by every simple closed geodesic; more generally, if ν is any measured lamination on R with support different than Λ, then the geometric intersection number satisfies $\iota(\mu, \nu) \neq 0$. Arational laminations Λ also have the property that every half-leaf is dense; in particular Λ is minimal. Uniquely ergodic laminations are arational.

(iv) The discussion works as well on compact surfaces with boundary. However the simple geodesics one works with are not allowed to be parallel to boundary components.

We summarize by listing the following adjectives attached to geodesic Λ or measured geodesic laminations (Λ, μ) on a finite area hyperbolic surface R:

arational measured lamination Each component of $R \setminus \Lambda$ is an ideal polygon possibly containing one puncture of R. Λ has positive intersection number with every closed geodesic; in particular it has no closed leaves. An arational measured lamination is minimal.

maximal or filling lamination Λ is not a proper subset of another lamination; each component of $R \setminus \Lambda$ is an ideal polygon, possibly containing a puncture.

minimal or connected lamination The support Λ has no sublaminations; either Λ consists of a single closed geodesic, or every leaf λ has infinite length and each half-leaf is dense in Λ. Every lamination is the union of finitely many minimal sublaminations and, if Λ is not measured, possibly finitely many isolated* leaves whose ends spiral in to the minimal laminations or end at a cusp.

filling or binding pair Two laminations Λ_1 and Λ_2 form a *filling pair*, and Λ_1 and Λ_2 *fill up* R, if every component of $R \setminus (\Lambda_1 \cup \Lambda_2)$ is a (simply connected) polygon or is a polygon containing a puncture of R. A filling pair satisfies $\iota(\gamma, \Lambda_1) + \iota(\gamma, \Lambda_2) > 0$ for every simple closed geodesic γ.

uniquely ergodic lamination There is one and only one measure μ with support Λ, up to positive multiples. The support of a uniquely ergodic measured lamination is minimal, but not necessarily maximal.

3.10 The convex hull of the limit set

Fenchel had long advocated using the convex core construction in $\mathbb{H}^3$ to study kleinian groups, since in his work with Nielsen he had found the corresponding construction in $\mathbb{H}^2$ for fuchsian groups very useful. However the difficulty was not in the construction, but in the analysis of the convex hull boundary. It was Thurston who taught us how to use the convex hull as an effective tool. The application required prior development of the theory of measured laminations.

In describing the theory, we will stick with the upper half-space model. We start with a closed set $\Lambda \subset \mathbb{C} \cup \infty \equiv \mathbb{S}^2$, with a nonempty complement $\Omega = \mathbb{S}^2 \setminus \Lambda$. The hyperbolic convex hull of Λ is defined as follows.

Let $C \subset \overline{\Omega}$ be a round circle in $\mathbb{S}^2$ that bounds an open disk $\Delta \subset \Omega$. If Λ is connected so that each component of Ω is simply connected, any circle in $\overline{\Omega}$ will determine such a disk. The circle C in turn determines a hyperbolic plane $C^* \in \mathbb{H}^3$. Denote by $H(C)$ the relatively closed half-space bounded by C^* that abuts the *exterior* of Δ. The (hyperbolic) *convex hull* of Λ is the relatively closed set

$$\widehat{\mathcal{C}}(\Lambda) = \bigcap_{C \subset \overline{\Omega}} H(C). \tag{3.7}$$

In constructing $\widehat{\mathcal{C}}(\Lambda)$ it suffices to restrict attention to maximal disks Δ — those that are not proper subsets of larger disks in Ω. The circle bounding a maximal disk meets $\partial\Omega$ in at least two points.

Since $\widehat{\mathcal{C}}(\Lambda)$ is convex, the (hyperbolic) line segment joining any two of its points lies in the set. In fact any geodesic with endpoints in $\partial\Omega = \Lambda$ is contained in $\widehat{\mathcal{C}}$. With

* A leaf λ is isolated if every point $p \in \lambda$ has a neighborhood U with $U \cap \lambda$ an arc through p.

Peter Storm one can define $\widehat{\mathcal{C}}(\Lambda)$ as the package obtained by shrink wrapping the set of all geodesics with endpoints in Λ.

The relative boundary $\partial\widehat{\mathcal{C}}(\Lambda) \subset \mathbb{H}^3$ is the union of *flat pieces* and *bending lines*.

A flat piece is a noncompact hyperbolic polygon contained in one of the hyperbolic planes C^* used to form the convex hull. It lies in the plane determined by a maximal disk that is bounded by a circle that meets $\partial\Omega$ in at least three points.

The complement in $\partial\widehat{C}(\Lambda)$ of the union of open flat pieces is the closed set of bending lines. A bending line ℓ is a geodesic whose endpoints lie in $\partial\Omega$. Distinct bending lines are disjoint but they possibly have a common end point. There are in general an uncountable number of them. The limit of a sequence of bending lines is either a bending line or a point in the common boundary $\partial\widehat{C} = \partial\Omega$. A flat piece, if not a whole plane, is bounded by bending lines.

An isolated bending line ℓ is the common boundary of adjacent flat pieces. The *bending angle* at ℓ is taken to be the exterior bending angle α so that $\alpha = 0$ corresponds to no bending at all and $\alpha = \pi$ corresponds to one flat piece folded over the other.

Each component S of the relative boundary $\partial\widehat{\mathcal{C}}(\Lambda) \cap \mathbb{H}^3$ faces a component Ω_S of Ω. It helps to keep in mind the picture of a domed stadium, such as one finds in Minneapolis. The *floor* of the stadium is Ω_S and the *dome* is S.

There is a continuous map $r : \Omega_S \to S$ called the *nearest point retraction*. This is defined as follows: Given $z \in \Omega_S$ examine the family of horospheres tangent to $\partial\mathbb{H}^3$ at z. This family depends on a parameter, for example the euclidean diameter. Exactly one of these spheres just touches S, necessarily at a single point, without crossing S. This point of first touching is called the *nearest point* and is denoted by $r(z)$. If $r(z)$ is in a flat piece, then there is a geodesic ray from $r(z)$, where it is orthogonal to $\partial\hat{\mathcal{C}}(\Lambda)$, ending at z. An isolated bending line $\ell \in S$ with bending angle α will be the image under r of a crescent $C_\ell \in \Omega_S$ with vertices in $\partial\Omega_S$.

The crescent C_ℓ is constructed as follows. There are two planes C_1^*, C_2^* rising from maximal circles C_1, C_2 and intersecting with exterior angle α. The angle interior to $\widehat{\mathcal{C}}(\Lambda)$ is $\pi - \alpha$. The sides of C_ℓ are orthogonal to C_1, C_2. Therefore the interior vertex angles of C_ℓ are

$$\alpha = 2\pi - \left(\frac{\pi}{2} + (\pi - \alpha) + \frac{\pi}{2}\right).$$

In particular, C_ℓ is not the crescent formed by $C_1 \cap C_2$ unless C_1 and C_2 are orthogonal.

Distinct isolated bending lines correspond to nonoverlapping crescents in Ω_S. If there are no isolated bending lines, r is a homeomorphism.

The nearest point retraction r fixes the points on the common boundary $\partial\Omega_S = \partial S$. Convex hulls are studied in detail in [Epstein et al. 2004].

Examples

In the degenerate case that Λ has exactly two points, the convex hull is simply the geodesic between the two points.

If Λ is the half infinite line $[0, +\infty]$ the convex hull is the vertical wall arising from the line. Like the above example, this is a degenerate case in that the interior of the convex hull is empty. However in this case, its boundary is regarded as the union of the two sides of the wall, with exterior bending angle π.

The dome over a round disk Δ, is the plane rising from the circle $\partial\Delta$. How about two round disks with angle of intersection α measured exterior to one disk and interior to the other? The dome over the union consists of two flat pieces meeting with exterior bending angle α. There are two flat pieces and one bending line.

The dome over the region bounded by an ellipse is a half-ellipsoid. There is a continuous family of bending lines which sweep out the dome which is a smooth surface. There are no flat pieces and the dome is a smooth, ruled surface.

Next consider a wedge $W = \{z \in \mathbb{C},\ 0 \le \arg z < \alpha \le \pi\}$. If $\Lambda = W$, the convex hull boundary consists of the two flat pieces rising from the edges of W and one bending line. The exterior bending angle is $\pi - \alpha$. If instead Λ is the closure of the complement of W, then the dome over W is a half cone. Again it is swept out by the bending lines; there are no flat pieces.

The dome over a convex euclidean triangle contains one flat piece which is contained in the plane rising from the maximal inscribed circle, and parts of three cones near the vertices. The dome is a smooth C^1-surface. In fact the dome over any euclidean convex region is a smooth surface [Epstein et al. 2006; 2004].

The bending measure

Each component $S = \mathrm{Dome}(\Omega_S)$ which is not a whole plane carries a nonzero *bending measure*. At an isolated bending line, it is just the atomic measure with support on the line given by the exterior bending angle. In general, the bending measure is constructed by a process akin to Riemann integration, that is, by approximating the dome by a sequence of finitely bent surfaces. The basic result is the following theorem of Thurston; the detailed proof appears in [Epstein and Marden 1987].

Theorem 3.10.1. *Suppose Ω is a simply connected region whose complement Λ in $\mathbb{S}^2$ has at least three points.*

(i) *The hyperbolic metric in $\mathbb{H}^3$ restricts to give a path metric on Dome(Ω) referred to as its hyperbolic metric.*
(ii) *There is an isometry in the respective hyperbolic metrics $\Upsilon : \mathrm{Dome}(\Omega) \to \mathbb{H}^2$.*
(iii) *Under Υ, the set of bending lines is carried to a geodesic lamination Λ in $\mathbb{H}^2$ and the bending measure on Dome(Ω) is carried to a (bounded) transverse measure on Λ.*
(iv) *If Ω is invariant under a kleinian group G, then Dome(Ω) and the set of bending lines are also G-invariant. The corresponding measured lamination in $\mathbb{H}^2$ is invariant under the fuchsian group $\Upsilon G \Upsilon^{-1}$.*

The hyperbolic metric on the simply connected region Ω is carried over from $\mathbb{H}^2$ by a Riemann mapping. In terms of the hyperbolic metrics on Ω and its dome, the nearest point retraction $r : \Omega \to \mathrm{Dome}(\Omega)$ satisfies $d(r(z_1), r(z_2)) \le 2d(z_1, z_2)$; that is, r is 2-Lipschitz [Epstein et al. 2004]. If Dome(Ω) is instead infinitely connected, one can pass to its universal cover and map that and its measured bending lamination to $\mathbb{H}^2$.

Now suppose $\Lambda(G)$ is the limit set of a kleinian group G. Its convex hull $\widehat{\mathcal{C}}(G)$ is G-invariant. Each relative boundary component S of $\widehat{\mathcal{C}}(G)$ is the dome over a component Ω_S of $\Omega(G)$ and is invariant under $\mathrm{Stab}(\Omega_S)$.

The convex hull $\widehat{\mathcal{C}}(G)$ necessarily contains the axes of all loxodromics of G since these have endpoints in the limit set. Can the axis of a $g \in G$ be a bending line? Only if the trace of g is real with $|\mathrm{tr}\, g| > 2$. Otherwise the angular part of the trace would force a rotation about the axis, and therefore could not preserve the convex hull.

The section cannot be closed without mentioning the following remarkable fact described by Dennis Sullivan. For a full discussion and proof see [Epstein and Marden 1987] or [Epstein et al. 2004].

Theorem 3.10.2 (Sullivan Convex Hull Theorem). *There exists a universal constant* $1 < K < 14$ *with the following property. Given any simply connected region* $\Omega \subset \mathbb{C}$, $\Omega \neq \mathbb{C}$, *there exists a* K*-quasiconformal mapping* $F : \Omega \to \mathrm{Dome}(\Omega)$ *which extends to pointwise fix every point on the common boundary* $\partial\Omega$.

If Ω *is invariant under a group* Γ *of Möbius transformations*, F *can be chosen to satisfy additionally* $F \circ \gamma = \gamma \circ F$ *for all* $\gamma \in \Gamma$.

Pleated surfaces

We have spoken of the structure of a convex hull boundary component, especially the dome over a simply connected region. Now consider the reverse process. That is, given a measured lamination (Λ, μ) in $\mathbb{H}^2$, can we construct a surface in $\mathbb{H}^3$ whose bending measure is μ?

Let's start with the simplest cases. Take the equatorial plane $\mathbb{H}^2$ (the unit disk) in the ball model and fix a diameter ℓ. Bend $\mathbb{H}^2$ along ℓ with exterior bending angle $0 < \theta < \pi$. Here $\theta = 0$ corresponds to no bending at all. The other extreme $\theta = \pi$ corresponds to two situations: (i) folding $\mathbb{H}^2$ in half along ℓ, or (more commonly) (ii) pushing ℓ out to ∞ to become a single point ξ thereby forcing $\mathbb{H}^2$ in the limit to become two hyperbolic planes whose boundaries are tangent at ξ so that one plane is the image of the other under a designated parabolic with fixed point ξ.

To normalize the direction of bending, bend so that the result lies in the upper half of the ball. The resulting "pleated surface" S bounds on one side a convex region whose floor is bounded by two circular arcs with interior bending angle $\pi - \theta$. The dome has only one bending line.

The construction is easily generalized to a finite system of ordered, mutually disjoint hyperbolic lines, possibly with common endpoints, $\ell_1, \dots, \ell_k \subset \mathbb{H}^2$. Assign an exterior bending angle $0 < \theta_i < \pi$ to each line. Then systematically bend the plane $\mathbb{H}^2$. For example we may assume that first bend along ℓ_1 results in $P_1 = P$ constructed above. Then in P_1 locate the copy of ℓ_2, say it lies to the right of ℓ_1. Then bend the

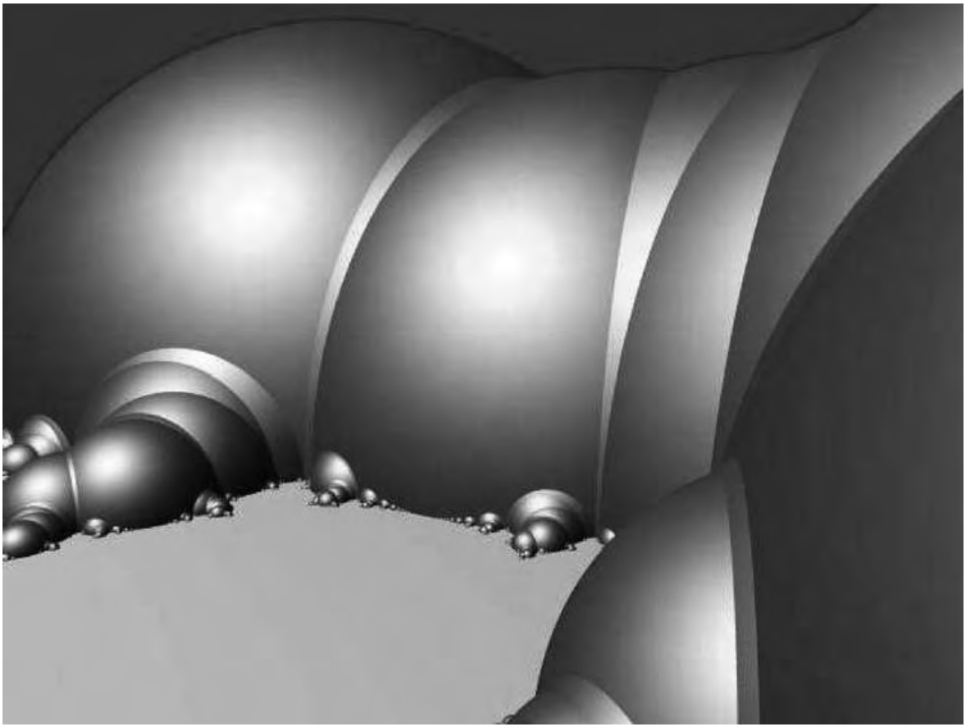


Fig. 3.7. A section of the dome over a component of a quasifuchsian ordinary set.

half-plane in P_1 lying to the right of ℓ_1 along ℓ_2 with exterior angle θ_2. And so on for all the lines. We end up with what is called a *pleated surface* P_k. It is locally convex but is not necessarily embedded in $\mathbb{H}^3$ — it may well have self-intersections. It has k bending lines, the images of the $\{\ell_i\}$. In any case there is a hyperbolic isometry $\Upsilon : \mathbb{H}^2 \to P_k$ — such that Υ^{-1} is just unbending. The finite measured lamination is carried to the bending lines and bending measure on P_k.

The same construction can be carried out given a general lamination Λ in $\mathbb{H}^2$ and a positive transverse Borel measure by using finite approximations. In fact it equally works for a real valued transverse Borel measure. In the general case the pleated surface has both positive and negative bending. It may not be locally embedded and may even be dense in all $\mathbb{H}^3$. The construction is such that if (Λ, μ) is invariant under a fuchsian group G, a deformation of G to a homomorphic image H is automatically determined. H is a group of Möbius transformations acting in $\mathbb{H}^3$ that map the pleated surface onto itself in a manner reflecting the action of G in $\mathbb{H}^2$, but H is unlikely to be discrete. The details are carried out in [Epstein and Marden 1987].

Another way of constructing a pleated surface from a geodesic lamination $\Lambda \subset \mathbb{H}^2$ is as follows. Suppose Λ is such that all gaps are ideal triangles; this is the generic case. If there is an injection $f : \partial\mathbb{H}^2 \to \partial\mathbb{H}^3$ (for example, the restriction of a quasiconformal deformation of a fuchsian group) then each leaf $\ell \subset \Lambda$ can be mapped to the line determined by the f-images of the endpoints of ℓ and the ideal triangles can then be filled in. If in addition Λ is invariant under a fuchsian group G and f conjugates G to a quasifuchsian group H then the lamination Λ/G gives rise to a pleated surface in $\mathbb{H}^3/H$. For an example, see Exercise 6-3.

Since the convex hull contains all geodesics, a flat piece of a pleated surface that is bounded by two or more geodesics lies in the convex hull. Thus most pleated surfaces lie entirely in the convex hull.

Formally, a *pleated surface* is determined by a *pleating map* $f : S \to \mathcal{M}$ of a hyperbolic surface S into a hyperbolic 3-manifold $\mathcal{M}$ with these properties:

(i) f takes any rectifiable path in S to a path in $\mathcal{M}$ of the same length.
(ii) Every $z \in S$ lies in an open geodesic arc which f maps to a geodesic arc in $\mathcal{M}$.
(iii) f sends cusps to cusps: it sends a small neighborhood of a cusp of S into a small neighborhood of a cusp of $\mathcal{M}$; the homomorphism $f_* : \pi_1(S) \to \pi_1(\mathcal{M})$ sends parabolics to parabolics.

Assumption (i) can replaced by (i'): geodesic paths in S are sent to rectifiable paths of the same length in $\mathcal{M}$. The apparently stronger definition is equivalent [Canary et al. 1987, II.5.2.6]. We may equally work with a lift of f to the universal covers.

The pleated surface is called *incompressible* if $f_* : \pi_1(S) \to \pi_1(\mathcal{M})$ is injective.

The *pleating locus* is the set $\Lambda \subset S$ consisting of those points $z \in S$ with the following property. There is one and only one open geodesic arc (up to inclusion) through z which f maps onto a geodesic arc in $\mathcal{M}$. The pleating locus Λ is a closed subset of S and is in fact a geodesic lamination. The image $f(\Lambda)$ is often referred to as the *pleating locus* as well, or as the *bending lines*. The map f is an isometry of the complementary gaps onto polygons in $\mathcal{M}$ that in general are infinitely sided.

Given such a general pleated surface, there is likely to be a great deal of positive and negative bending. Yet by associating a transverse segment τ to the set of positive endpoints on $\partial\mathbb{H}^3$ of the oriented leaves through τ and then a continuum in $\partial\mathbb{H}^3$, it is possible to construct a kind of bending measure which however is only finitely additive. This measure and the pleating locus characterize the pleated surface. For the details see [Bonahon 2001; 1996].

Given a lamination $\Lambda \subset S$ and a hyperbolic manifold $\mathcal{M}$, the lamination Λ is said to be *realizable* in $\mathcal{M}$ if there is a pleating map $f : S \to \mathcal{M}$ whose pleating locus contains Λ.

Suppose Λ is a finite geodesic lamination on S and $h : S \to \mathcal{M}$ is a map such that $h_* : \pi_1(S) \to \pi_1(\mathcal{M})$ is injective and cusps correspond to cusps. Assume in addition h is homotopic to a map $h' : S \to \mathcal{M}$ whose restriction to each leaf ℓ of Λ is a homeomorphism of ℓ onto a geodesic of $\mathcal{M}$. Then the conformal structure on S can be changed so that in the new structure, h is homotopic to a pleating map into $\mathcal{M}$, with pleating locus Λ. If Λ is maximal lamination and $\mathcal{M}$ is geometrically finite, the new hyperbolic structure needed on S is uniquely determined [Canary et al. 1987, II.5.3.11]. In these results, S need only be a surface of finite topological type.

However in normal practice, S is always a finite area surface. Such maps h will arise when we take earthquakes (Exercise 3-32) followed by bending. In this case the bending determines a pleated surface on the new hyperbolic structure resulting from the earthquake.

Proposition 3.10.3 [Thurston 1986b, Proposition 5.3]. *Given an $\varepsilon > 0$ that determines the thick/thin decomposition of $\mathcal{M}$ and given a constant $A > 0$, there exists $C > 0$ with the following property. Any incompressible pleated surface $f : S \to \mathcal{M}$ with* $\mathrm{Area}(S) \le A$ *satisfies*

$$\mathrm{Inj}_{\mathcal{M}}(f(x)) \le \mathrm{Inj}_S(x) \le C\,\mathrm{Inj}_{\mathcal{M}}(f(x)),$$

provided the distance of $f(x)$ from any closed geodesic in $\mathcal{M}$ of length not exceeding ε is at least 1.

That the injectivity radius is $r = \mathrm{Inj}_{\mathcal{M}}(f(x))$ means there is a hyperbolic ball in $\mathcal{M}$ of radius r, centered at $f(x)$, whose interior is embedded in $\mathcal{M}$, and no larger ball has this property. The uniform injectivity property guarantees that the injectivity radius in $\mathcal{M}$ at $f(x)$ is not substantially different from the injectivity radius on S at x, provided that $f(x)$ is not too close to a short geodesic in $\mathcal{M}$. The proof uses the fact that there is an upper bound for the injectivity radii on S in terms of A.

Consider now for simplicity the case of a closed hyperbolic surface S. The pleated surface $f : S \to f(S) \subset \mathcal{M}$ is called *doubly incompressible* if, in addition to being incompressible, (i) two loops on $f(S)$ which are freely homotopic in $\mathcal{M}$ come from loops which are already freely homotopic in S, and (ii) under f_*, maximal cyclic subgroups of $\pi_1(S)$ are sent to maximal subgroups of $\pi_1(\mathcal{M})$ (primitive elements are preserved). There is an important injectivity property for such pleated surfaces as follows (see [Minsky 2000] for the statement when there are parabolics and the application to the proof of the ending lamination conjecture).

Theorem 3.10.4 (Uniform injectivity of pleated surfaces ([Thurston 1986b, Theorem 5.2])). *Fix a closed hyperbolic surface S and a constant $\epsilon_* > 0$. Given $\varepsilon > 0$ there exists $\delta > 0$ such that the following property holds for any doubly incompressible pleated surface $f : S \to \mathcal{M}$: Let $\Lambda \subset S$ denote the lamination representing the pleating locus $f(\Lambda)$. Then if $x, y \in \Lambda$ are in the ϵ_*-thick part of S,*

$$d_{T(\mathcal{M})}(v_x, v_y) \le \delta \quad \text{implies} \quad d(x, y) \le \varepsilon.$$

Here v_x and v_y denote the unit tangent vectors to the leaves of $f(\Lambda)$ containing $f(x)$ and $f(y)$, respectively. Their distance apart is measured in the projectivized tangent bundle $T(\mathcal{M})$. The theorem says that when the tangent vectors are not too far from being parallel, the initial points x, y (and the leaves of Λ containing them) are ε-close in S. In particular the bit of $f(S)$ bounded by the lines containing $f(x)$, $f(y)$ is not wildly oscillating.

It is tough to give conditions on a measured lamination (Λ, μ) with μ a positive measure, so that the corresponding pleated surface is the dome over a simply connected region. The best result is in terms of the norm $\|\mu\| = \sup_\sigma \mu(\sigma)$, where σ ranges over all transverse segments of unit length. In [Epstein et al. 2004] it is shown that there exists a constant $0 < c \le 2 \arcsin \tanh(\frac{1}{2}) \cong 0.96$, with the following property. If $\|\mu\| < c$ then (Λ, μ) is the bending measure of $\mathrm{Dome}(\Omega)$ for some simply connected region Ω. It is conjectured that the upper bound given is best possible; in any case it is known that it cannot be larger. On the other hand it is known that if the pleated surface is a dome, then $\|\mu\| < 4.88$ [Bridgeman 2003].

3.11 The convex core

The quotient

$$\widehat{\mathcal{C}}(\Lambda)/G = \mathcal{C}(G) \subset \mathcal{M}(G)^{\text{int}}$$

is hyperbolically convex and is called the *convex core of* $\mathcal{M}(G)$. Every closed geodesic in $\mathcal{M}(G)$ lies in $\mathcal{C}(G)$. Indeed the convex core can be defined to be the smallest convex set with this property. The inclusion $\mathcal{C} \hookrightarrow \mathcal{M}(G)$ is an isomorphism between fundamental groups. Thus the convex core is representative of the full manifold. At one extreme, for fuchsian groups the convex core is flat without interior. At the other extreme, if $\Omega(G)$ is empty then the convex core is the full manifold $\mathcal{M}(G)$.

Here are two additional facts about convex cores.

(1) *The nearest point retraction projects to the quotient and is a continuous map from each component of* $\partial\mathcal{M}(G)$ *to the component of* $\partial\mathcal{C}(G)$ *that it faces.*
(2) *If* G *is not fuchsian, then* G *is geometrically finite if and only if* $\mathcal{C}(G)$ *has finite volume.*

The reason fuchsian groups are excluded is that their three-dimensional convex core always has zero volume, even if the group is not finitely generated. The exclusion could be avoided by requiring that an ε-neighborhood of the core be of finite volume. Instead we will simply exclude fuchsian groups from the statement. For fuchsian groups, Fenchel and Nielsen made good use of the fact that the group is finitely generated if and only if its convex core with respect to $\mathbb{H}^2$ has finite area.

Proof of item (2). We begin with a lemma.

Lemma 3.11.1. *Suppose* G *is nonelementary, has no elliptics and* $\zeta \in \partial\mathbb{H}^3$ *is a parabolic fixed point.*

(i) *If* Stab_ζ *is a rank two parabolic group, then* $\widehat{\mathcal{C}}(\Lambda)$ *contains a horoball* $\mathcal{H}_\zeta$ *at* ζ.
(ii) *Suppose* Stab_ζ *is rank one and supports a double horocycle* σ_1, σ_2 *at* ζ. *For some horoball* $\mathcal{H}_\zeta$, $\widehat{\mathcal{C}}(\Lambda)$ *contains* $\mathcal{H}_\zeta \cap (H_1 \cap H_2)$. *Here* H_i *is the half-space, bounded by the plane rising from the horocycle* $\sigma_i = \partial\Delta_i$, *which abuts the exterior of its horodisk* Δ_i.

Proof. We may assume that in the upper half-space model $\zeta = \infty$ and $T_1(z) = z + 1$ is a generator of Stab_ζ.

In the rank two case, there is another generator $T_2(z) = z + a$ and we may assume that $|a| \geq 1$. We claim there is a maximal diameter d for circles $C \subset \overline{\Omega(G)}$ that bound disks in $\Omega(G)$. Suppose otherwise. When C has a sufficiently large diameter the disk Δ that it bounds will have the property that $\Delta^* = \Delta \cup T_1(\Delta) \cup T_2(\Delta) \cup T_1 T_2(\Delta)$ is simply connected. But then the orbit of Δ^* under Stab_ζ covers $\mathbb{C}$. This means that $\Omega(G) = \mathbb{C}$ which is impossible if G is not Stab_ζ itself. Consequently the horoball $\{(z, t) : t > d/2\}$ is contained in $\widehat{C}(\Lambda)$.

In the rank-one case, because there is a double horodisk at ∞, $\Lambda(G)$ is contained inside a minimal width strip $\mathcal{S} = \{z : b_1 \leq \operatorname{Im} z \leq b_2\}$ where both horizontal lines $\operatorname{Im} z = b_1, b_2$ contain limit points. In fact their intersection with the vertical strip $\mathcal{V} = \{z : 0 \leq \operatorname{Re} z \leq 1\}$ also contains limit points. We see that there is a maximal diameter $d < \infty$ for circles $C \subset \overline{\Omega(G)}$ centered in $\mathcal{V}$ that bound disks in $\Omega(G)$. Now the maximal horocycles σ_1, σ_2 are bounded by $\operatorname{Im} z = b_1, b_2$. These two observations translate into the second statement. $\square$

We continue our proof of item (2). If $\mathcal{M}(G)$ is geometrically finite, parallel to each component of $\partial\mathcal{M}(G)$ is a component of $\partial\mathcal{C}(G)$. Each cusp is taken care of by Lemma 3.11.1. The convex hull has finite volume.

Conversely assume the convex hull has finite volume. There are at most a finite number of solid cusp tori in $\mathcal{C}(G)$, according to Lemma 3.11.1. Ahlfors' Finiteness Theorem implies that $\mathcal{M}(G)$ has a finite number of boundary components and each is a closed surface with at most a finite number of punctures. Each component is parallel to a boundary component of the convex hull with the same property. Consequently $\mathcal{C}(G)$ has a finite number of boundary components.

Let $\mathcal{C}_\varepsilon$ denote the ε-neighborhood of $\mathcal{C}(G)$ in $\mathcal{M}(G)$ — the set of points of distance $< \varepsilon$ from the convex hull. The volume of $\mathcal{C}_\varepsilon$ is also finite because each ε-ball is either contained in the interior or it intersects a boundary component. The thick part of the core is covered by a finite number of these ε-balls and is therefore compact.

The boundary of the thick part of the core contains a compact piece of $\partial\mathcal{C}(G)$, the boundaries of tubes about short geodesics, pairing cylinders, and cusp tori. The core cannot contain entire cusp cylinders — the projection of horospheres at rank one cusps — because these are not compact. Instead it is their intersections with $\mathcal{C}(G)$ that are compact. We conclude that each truncated cusp cylinder pairs two punctures on $\partial\mathcal{C}(G)$, and there are a finite number of them. So $\mathcal{M}(G)$ itself has the essential compactness of a geometrically finite manifold. $\square$

Length estimates

Suppose $\widehat{S}$ is a simply connected boundary component of $\widehat{\mathcal{C}}(\Gamma)$, and S is the corresponding boundary component of the convex core $\mathcal{C}(G)$ of $\mathcal{M}(G)$. The surface S faces a component R of $\partial\mathcal{M}(G)$. If γ_S is a closed geodesic in S, there is a uniquely determined geodesic γ_R in the hyperbolic metric on R which is freely homotopic to γ_S. In the following theorem, the lower bound is obtained from the best current estimate for the equivariant "K" in Theorem 3.10.2 and the fact that the minimal

Lipschitz constant in the same homotopy class does not exceed this “K”. The upper bound follows directly from the fact that the Lipschitz constant of the nearest point retraction does not exceed 2. The inequality shows that the hyperbolic geometry of the two surfaces is tightly bound together.

Theorem 3.11.2 [Epstein et al. 2004]. *In the respective hyperbolic metrics,*

$$\frac{1}{14} < \frac{\ell(\gamma_S)}{\ell(\gamma_R)} \leq 2.$$

The same bounds hold for the lengths of corresponding measured laminations.

The length of the geodesic γ_M in the interior of $\mathcal{M}(G)$ freely homotopic to γ_R is likewise bounded by $2\ell(\gamma_R)$ as shown in Exercise 5-2 — γ_M will be identical to γ_S if γ_S is a bending line.

According to [Bridgeman 1998], there exists a universal constant B with the following property. If S is a component of $\partial\mathcal{C}(G)$ as above, then

$$\ell(\beta_S) \leq B\pi^2|\chi(S)|, \tag{3.8}$$

where $\ell(\beta_S)$ is the length on S of the bending lamination β_S and $\chi(S)$ is the Euler characteristic. In particular if β_S is supported on a single geodesic of length L_β with bending angle θ, $\ell(\beta_S) = L_\beta \cdot \theta$ is bounded; the longer L_β is, the smaller θ must be.

Existence of bending measures

There is a beautiful recent result of Bonahon and Otal [2004], completed by Lecuire [2003], characterizing geometrically finite groups by the bending laminations of their convex core boundaries. See [Lecuire 2004a; 2004b] for further applications of this subject.

Start with an orientable, compact manifold M^3 other than a solid torus T^2 or a thickened torus $T^2 \times [0, 1]$, and whose interior has a hyperbolic structure. Thus M^3 is a model for a geometrically finite manifold with solid cusp tubes and cusp tori removed. We assume that ∂M^3 has some nontorus components, which may or may not be incompressible, and we may as well assume each has a hyperbolic structure. Let (Λ, μ) be a measured lamination on the nontorus components of ∂M^3. We allow that on some boundary components, (Λ, μ) may be the zero lamination (no leaves).

On a closed leaf γ of Λ, μ has atomic measure $\mu(\gamma) > 0$ which we will think of as a bending angle. Let D and C be an essential disk and cylinder in M^3. As we know, the *geometric intersection number* $\iota(\partial D, \Lambda)$ or $\iota(\partial C, \Lambda)$ is the generalization of the case that Λ consists of a finite number of closed leaves and μ is the unit atomic measure on each. In the finite case, $\iota(\partial D, \Lambda)$ or $\iota(\partial C, \Lambda)$ the minimum number of times that simple loops freely homotopic to ∂D cross the leaves of Λ or the minimum number of times simple loops freely homotopic to the components of ∂C cross the leaves of Λ. We are assuming that Λ has at least one leaf, yet it is possible that one or more nontorus components of $\partial\mathcal{M}$ carry no leaves.

Theorem 3.11.3 (Existence of bending measures [Bonahon and Otal 2004; Lecuire 2003]). *Given the measured lamination (Λ, μ) on ∂M^3, there exists a geometrically finite, nonfuchsian, $\mathcal{M}(G_\mu)$ whose convex core boundary has the bending lamination (Λ, μ) if and only if the following conditions are satisfied*:

(*i*) *On each closed leaf α, $\mu(\alpha)$ satisfies $0 < \mu(\alpha) \le \pi$.*
(*ii*) *For each essential disk $D \subset M^3$, $\iota(\partial D, \Lambda) > 2\pi$.*
(*iii*) *There exists $\eta > 0$ such that $\iota(\partial C, \Lambda) \ge \eta$ for each essential cylinder $C \subset M^3$.*

If Λ consists of a finite number of closed leaves then the kleinian group G_μ is uniquely determined up to Möbius equivalence.

The closed leaves γ with $\mu(\gamma) = \pi$ will correspond to the rank one cusps of $\mathcal{M}(G_\mu)$. Of course the torus boundary components of ∂M^3 will correspond to the rank two cusps of $\mathcal{M}(G_\mu)$. The proof of uniqueness is outlined in Exercise 6-3. Uniqueness for all laminations is known for the once-punctured torus quasifuchsian case [Series 2004] and conjectured for the general case.

Suppose in addition that M^3 is compact, boundary incompressible, and has no essential cylinders (see Exercise 3-17). Assume we are given a maximal finite lamination of $\sum(3g_i - 3)$ simple closed geodesics $\Lambda = \cup\beta_j$ on $\partial\mathcal{M}^3$, g_i the genus of the i-th component of ∂M^3, and an atomic measure $0 < \mu(\beta_j) < \pi$ for each index. According to Theorem 3.11.3, there exists a uniquely determined (up to isometry) hyperbolic structure $\mathcal{M}(G_\mu)$ on M^3 whose convex core $\mathcal{C}(G)$ has exactly the bending lamination (Λ, μ). In [Choi and Series 2006] it is shown that the $\sum(3g_i-3)$-complex lengths in $\mathcal{M}(G_\mu)$ (see Section 7.4) of the geodesics $\{\beta_j\}$ serve as local coordinates for the local deformations of $\mathcal{M}(G_\mu)$ in the representation variety $\mathfrak{R}(G_\mu)$ (see Section 5.1).

If the lamination is finite, condition (ii) on the geometric intersection number ι requires that the boundary of each essential disk has at least three essential crossings with Λ. Condition (iii) insures that if one boundary curve of C is a leaf of Λ then the other must be transverse to Λ.

If a nontorus component of ∂M^3 carries no leaves, Theorem 3.11.3 provides that the corresponding component R of $\partial\mathcal{M}(G_\mu)$ is *totally geodesic*. This means that every component Ω_R of $\Omega(G_\mu)$ lying over R is a round disk; the convex hull boundary component that faces R is a hyperbolic plane. A compressible boundary component cannot become totally geodesic; in line with this fact condition (ii) requires compressible components to contain leaves of Λ.

To understand why condition (ii) is necessary, suppose we have a compact convex core with bending lines as simple loops $\{\sigma_i\}$ with exterior bending angles $\{\beta_i\}$. Consider a compressing disk D with $\partial D \subset \partial\mathcal{C}(G)$. We may assume ∂D is piecewise geodesic and D is piecewise flat. The Gauss–Bonnet formula (1.3) tells us that

$$\sum \beta_i \iota(\partial D, \sigma_i) = \text{Area}(D) + 2\pi > 2\pi.$$

For (iii) suppose C with $\partial C \subset \partial \mathcal{C}(G)$ is an essential cylinder, also piecewise geodesic. We find that $\sum \beta_i \iota(\partial C, \sigma_i) = \text{Area}(C)$ so that we must have $\iota\big(C, \bigcup \sigma_i\big) > 0$.

The proof in the finite case starts by showing that there exists a geometrically finite $\mathcal{M}(G)$ homeomorphic to M^3 whose convex core is bent along $\sigma = \bigcup \sigma_i$. Then the manifold and bending angles are continuously deformed until they match the assigned angles. To establish existence the following argument is used. Remove half-tubular neighborhoods of the $\{\sigma_i\}$ and double the resulting manifold. This gives a compact manifold with tori boundary components. Make assumptions on $\{\sigma_i\}$ so the manifold is irreducible and atoroidal. As a consequence it has a complete hyperbolic structure of finite volume. One then uses the theory of cone manifolds (Exercises 4-7 and 6-3) to deform the rank two cusps to get a symmetric cone manifold with small cone angles. Undoubling results in the required convex hull.

A typical application is the following. Consider quasifuchsian groups representing a pair of surfaces of genus 2, say. For Λ, take a simple loop γ_{bot} on the "bottom" component and a finite number of mutually disjoint, nonparallel, simple loops $\{\beta_i\}$ on the top. To fulfill condition (iii) of Theorem 3.11.3 we must assume that *every* β_i is freely homotopic to a loop on the bottom component which is transverse to γ_{bot}. Assign positive atomic measures each less than π to all the simple loops. According to Theorem 3.11.3 there is a unique quasifuchsian group representing genus 2 surfaces whose convex hull boundary has the prescribed bending measure. By varying the measure on γ_{bot} while leaving the measures on $\{\beta_i\}$ fixed, we obtain a "slice" of the deformation space.

Parker and Series [1995] have an explicit construction for bending along one geodesic in the case of once-punctured torus quasifuchsian groups; see their bending formulas (8.39), (8.41).

3.12 The compact and relative compact core

There is another important "core" in a hyperbolic manifold, and this one is always compact but is not hyperbolic. It was discovered by Peter Scott [1973a] and independently by Peter Shalen:

In the interior of any hyperbolic manifold $\mathcal{M}(G)$ with G finitely generated there is a compact, connected, submanifold $C = C(G)$ such that (i) *inclusion of the fundamental group $\pi_1(C) \hookrightarrow \pi_1(\mathcal{M})$ is an isomorphism, and* (ii) *each component of ∂C bounds a noncompact component of* $\text{Int}(\mathcal{M}(G)) \setminus C$.

Property (ii) follows from (i). For if a complementary component were bounded by two components S_1, S_2 of ∂C, there would exist a simple loop in $\mathcal{M}(G)$ that crossed each of S_1, S_2 exactly once. Such a loop cannot be homotopic to a curve within C.

If $\mathcal{M}(G)$ is geometrically finite without parabolics, each component of ∂C is parallel to a component of $\partial \mathcal{M}(G)$. In the general case of no parabolics, each complementary component E of the core C is a neighborhood of exactly one *end* (see §5.5) of $\mathcal{M}(G)$.

The submanifold C is called a *compact core* of $\mathcal{M}(G)$. A core is uniquely determined up to homeomorphism: Two cores C_1, C_2 of $\mathcal{M}(G)$ are homeomorphic [McCullough et al. 1985]. An immediate consequence of its existence is that $\pi_1(\mathcal{M})$ is finitely presented, as stated in Theorem 2.5.3. Cores are a fundamental structure in studying geometrically infinite manifolds. According to [Bonahon 1986] (see also Exercise 3-11), each core can be cut along incompressible surfaces to result in a finite union of compression bodies and submanifolds with incompressible boundaries.

When there are parabolics, there is a useful refinement that incorporates the cusps. Namely, McCullough [1986] chooses a system of mutually disjoint horoballs in $\mathbb{H}^3$, associated with the parabolic fixed points, having the property that the union $\mathcal{H}$ is invariant under the action of G and the "parabolic locus" $\mathcal{P} = \mathcal{H}/G$ is embedded in $\mathcal{M}(G)$. The components of $\mathcal{P}$ are solid cusp tubes and solid cusp tori. Then $\mathcal{M}_p = \mathcal{M}(G)\setminus\mathcal{P}$ has the property that each component of the relative boundary $\partial\mathcal{M}_p$ is either a component of $\mathcal{M}(G)$, a cusp cylinder, or a cusp torus.

There exists a compact, connected, submanifold $C_{\mathrm{rel}} \subset \mathcal{M}_p \cap \mathrm{Int}(\mathcal{M}(G))$ *such that*

(*i*) *The inclusion* $\pi_1(C_{\mathrm{rel}}) \hookrightarrow \pi_1(\mathcal{M}(G))$ *is an isomorphism*,
(*ii*) *Each torus component of* $\partial\mathcal{P}$ *is a component of* $\partial C_{\mathrm{rel}}$,
(*iii*) *Each cylinder component of* $\partial\mathcal{P}$ *intersects* $\partial C_{\mathrm{rel}}$ *in a closed annular region*, *and*
(*iv*) *Each component of* $\partial C_{\mathrm{rel}} \setminus C_{\mathrm{rel}} \cap \partial\mathcal{P}$ *is the boundary of a noncompact component of* $\mathrm{Int}(\mathcal{M}_p) \setminus C_{\mathrm{rel}}$.

The submanifold C_{rel} is called a *relative compact core*.

If $\mathcal{M}(G)$ is geometrically finite, each component of $\partial C_{\mathrm{rel}} \setminus C_{\mathrm{rel}} \cap \partial\mathcal{P}$ is parallel to a component of $\partial\mathcal{M}(G)$.

Note that in the presence of rank one parabolics, $\partial\mathcal{M}(G)$ might be incompressible at the same time the boundary of the relative core is compressible. A simple example is a fuchsian group G representing a surface with punctures. When the interior of the solid pairing tubes are removed from $\mathcal{M}(G)$, the result is a handlebody. The compact core and relative compact cores are also handlebodies, but the relative core incorporates information about the punctures.

3.13 Rigidity

As mentioned in Chapter 1, hyperbolic polygons or convex polyhedra tend to be rigid — uniquely determined up to isometry by their angles. In dimensions larger than two, the same is true of finite volume hyperbolic manifolds.* Yet finite area surfaces are not rigid, except for the thrice punctured sphere. For this reason it caused a big stir in the kleinian world when Mostow first came up with his rigidity theorem for closed, hyperbolic n-manifolds, later extended to cusped manifolds in [Marden 1974a] (for $n = 3$), and independently in [Prasad 1973] for n dimensions. This was before the Thurston era, when we knew a lot less than we thought we did.

* In fact $\mathbb{H}^3$ itself is rigid in the sense that there is no nonconstant harmonic map of $\mathbb{H}^3$ into any riemannian 3-manifold of nonpositive sectional curvature [Leung and Wan 2001].

Mostow Rigidity Theorem [Mostow 1973]. *Suppose we have a hyperbolic manifold or orbifold $\mathcal{M}(G)$ of finite volume, with dimension $n \geq 3$, and an isomorphism $\phi : G \to H$ onto another kleinian group H. Then ϕ is determined by an isometry $\mathcal{M}(G) \to \mathcal{M}(H)$.*

In other words, $\mathcal{M}(G)$ is uniquely determined in the isomorphism class of G, up to orientation preserving or reversing Möbius equivalence. As already pointed out, rigidity does not hold in the hyperbolic plane: all surfaces with the same genus and the same number of punctures have the same (finite) area but are not usually isometric to each other. Surprisingly, the reason behind this state of affairs is that the limit set of a fuchsian group of finite area is $\mathbb{S}^1$, while for $n \geq 3$ the limit set is $\mathbb{S}^{n-1}$.

There is a homotopy analogue of Mostow's theorem as follows. This is a deep result by Gabai, Meyerhoff and N. Thurston that required a sophisticated computer program to complete. Later we will restate and discuss this theorem from a different point of view (page 247).

Theorem 3.13.1 [Gabai et al. 2003]. *Suppose M^3 is a closed, irreducible manifold* (for this notion see Section 6.3) *and $\mathcal{M}(G)$ a closed hyperbolic manifold. Assume there is an isomorphism $\phi : \pi_1(M^3) \to \pi_1(\mathcal{M}(G)) \cong G$. Then ϕ is induced by a homeomorphism $\Phi : M^3 \to \mathcal{M}(G)$.*

If Φ_1, Φ_2 are homotopic homeomorphisms, then Φ_1 is isotopic to Φ_2.

Of course if M^3 is also hyperbolic, the first statement is just Mostow's theorem with Φ an isometry. Even in this case the second statement is new. A homotopy is a continuous map $F : M^3 \times [0, 1] \to \mathcal{M}(G)$ such that $F(\cdot, 0) = \Phi_1$, $F(\cdot, 1) = \Phi_2$. For it to be an isotopy, each intermediate map $F(\cdot, t)$ must also be a homeomorphism. For example, a homotopy can send a geodesic α to a simple loop α' whose intersection with a tiny ball is knotted there. An isotopy cannot cause this effect. There is a subtle but important distinction between homotopy and isotopy. See Exercise 3-24.

For a further discussion of topological rigidity see Section 5.3.

Corollary 3.13.2. *If $\mathcal{M}(G)$ has finite volume, every orientation preserving (and orientation reversing) homeomorphism σ of $\mathcal{M}(G)$ onto itself is homotopic (isotopic if $\mathcal{M}(G)$ is closed) to an isometry.*

Unlike the case for finite area surfaces, the mapping class group of a finite volume hyperbolic 3-manifold is finite! We are implicitly assuming that a homotopy class contains at most one isometry, see Exercise 3-25.

In our study of quasifuchsian manifolds we have already made use of an analogue of Mostow rigidity for manifolds with boundary; in recent literature this is referred to as Marden's isomorphism (or rigidity) theorem [1974a]. See also [Tukia 1985b, Theorems 4.2, 4.7].

Theorem 3.13.3. *Suppose G is a geometrically finite group without elliptics and $\Phi : \Omega(G) \to \Omega(H)$ is a conformal mapping that induces an isomorphism $\phi : G \to H$*

by the correspondence $\phi(g)(z) = \Phi \circ g \circ \Phi^{-1}(z)$, $z \in \Omega(H)$. Then Φ is the restriction to $\Omega(G)$ of a Möbius transformation A and $\phi(g) = AgA^{-1}$ for all $g \in G$.

Thus $\mathcal{M}(G)$ is uniquely determined up to isometry by the isomorphism type of its fundamental group and the conformal structure of its boundary. In a sense, Mostow's theorem is a special case.

Sullivan [1981] established the most general form of rigidity that does not require geometric finiteness at all:

Sullivan Rigidity Theorem. *Suppose $\mu(z)$, $z \in \mathbb{S}^2$, is a Beltrami differential with respect to a finitely generated kleinian group G. Then $\mu(z) = 0$ for* a.e. $z \in \Lambda(G)$.

Corollary 3.13.4. *Geometrically infinite* (*as well as geometrically finite*) *manifolds $\mathcal{M}(G)$ with $\partial\mathcal{M}(G) = \varnothing$ are rigid under quasiconformal deformation.*

This result does not require any knowledge of the area of the limit set, nor does it give any information about its area. It says that the limit set can only support the zero Beltrami differential so that from the point of view of quasiconformal deformations, its area has no consequence. In the case that $\Omega(G) \neq \varnothing$, we now know that $\Lambda(G)$ has zero area — Ahlfors' conjecture is confirmed! (See Section 5.5.1.) So Sullivan's theorem for this case follows from the fact that a Beltrami differential needs only to be defined up to a set of zero measure for the Beltrami equation to have a solution (Section 2.8), uniquely determined up to postcomposition with a Möbius transformation. On the other hand when $\Lambda(G) = \mathbb{S}^2$, Sullivan's theorem still comes to the fore:

If $f : \mathbb{S}^2 \to \mathbb{S}^2$ is quasiconformal, induces an isomorphism $\phi : G \to H$, and, if $\Omega(G) \neq \varnothing$, restricts to a conformal map $\Omega(G) \to \Omega(H)$, then f is a Möbius transformation.

Outline of the proof of the Mostow Rigidity Theorem. The theorem holds in n-dimensional hyperbolic space but here we will stick to three.

The orbifold case of the theorem can be reduced to the manifold case by Selberg's lemma. Namely, given G, there is a torsion-free, normal subgroup $H \subset G$ of finite index; thus $\mathcal{M}(H)$ has finite volume, being finite-sheeted over $\mathcal{M}(G)$. The isomorphism ϕ restricts to the isomorphism $H \to \phi(H) = H_1 \subset \phi(G) = G_1$. Assuming the manifold case, $\phi : H \to H_1$ is a conjugation $h \in H \mapsto AhA^{-1}$, where we can assume A to be Möbius, rather than anti-Möbius, by replacing H if necessary with the group $H' = JHJ$, where J is a reflection in some hyperbolic plane. (The isomorphism $G \to H'$ is then given by $\phi(g) = AJgJA^{-1}$, for $g \in G$.)

Now the deck transformations form a finite group C of isometries of $\mathcal{M}(H)$ and likewise the deck transformations of $\mathcal{M}(H_1)$ over $\mathcal{M}(G_1)$. We conclude that G_1 itself is A-conjugate to G.

We will base our argument on the following beautiful theorem of Tukia:

Theorem 3.13.5 [Tukia 1985a]. *Suppose G is any nonelementary kleinian group, $\zeta \in \Lambda(G)$ is a conical limit point, and $f : \mathbb{S}^2 \to \mathbb{S}^2$ is a homeomorphism which is* *differentiable at ζ with nonzero derivative. Assume $\phi : G \to H$ is a homomorphism to another kleinian group given by $f \circ g(z) = \phi(g) \circ f(z)$ for all $g \in G,\ z \in \mathbb{S}^2$. Then f is a Möbius transformation*!

Proof. We will give the proof reported in [Kapovich 2001, Theorem 8.34]. We will assume the homeomorphism is orientation preserving, although this is not necessary.

The limit point $\zeta \in \Lambda(G)$ is a *conical limit point* or *point of approximation* if it has the following property. Let $\gamma(t)$, $0 \le t < \infty$, be a geodesic ray ending at ζ. Given a point $O \in \mathbb{H}^3$, there exists $r > 0$ such that there is an infinite subsequence of the orbit $G(O)$ that lies in the r-tubular neighborhood about γ (and hence converges to ζ). In the quotient manifold, the condition means that the projection of the ray $\gamma(t)$ is *recurrent* in the sense that it meets a ball of radius r about the projection of O for a sequence $\{t_n\}$, $t_n \to \infty$. A loxodromic fixed point is always a conical limit point but a parabolic fixed point is not. Beardon and Maskit [1974] proved that a kleinian group is geometrically finite if and only if all limit points, except parabolic fixed points, are conical limit points; see Exercise 3-18.

We may assume that $\zeta = 0 = f(0)$ and that O lies on the vertical axis rising from $z = 0$ in the upper half-space model. Let γ be the vertical segment descending from $O \in \mathbb{H}^3$ to $z = 0$. There is an infinite sequence $g_n \in G$ such that for some $r > 0$ and each large index n, the (hyperbolic) distance $d(g_n(O), \gamma) < r$. Find the point $y_n \in \gamma$ that is closest to $g_n(O)$; it is within distance r. Then find $a_n > 0$ such that the transformation $A_n : \vec{x} \mapsto a_n\vec{x}$ takes O to y_n; further, $\lim a_n = 0$. Passing to a subsequence if necessary we may also assume that $\lim g_n^{-1}A_n = B$ exists as a Möbius transformation (because the distance of $g_n^{-1}A_n(O)$ to O is uniformly bounded by r). Set

$$f_n(z) \ = \ a_n{}^{-1} f(a_n z) \ = \ A_n^{-1} \circ f \circ A_n(z), \quad z \in \mathbb{C}.$$

That the complex valued function $f(z)$ is differentiable at $z = 0$ with nonzero derivative means that there is a linear transformation $L : \mathbb{R}^2 \to \mathbb{R}^2$ (i.e., a 2×2 real matrix operating on $z \in \mathbb{C}$ as a vector), with nonzero determinant, such that

$$f(\Delta z) \ = \ L(\Delta z) + \epsilon(\Delta z)\Delta z, \qquad \lim_{\Delta z \to 0} \epsilon(\Delta z) = 0.$$

(Alternatively $L(z, \bar{z}) = az + b\bar{z}$, for some $a, b \in C$, $|a|^2 - |b|^2 > 0$.) Treating A_n as a linear transformation on vectors $z \in \mathbb{R}^2$ and setting $\Delta z = A_n(z)$, we obtain for f_n that

$$f_n(z) \ = \ L(z) + \epsilon(A_n(z))z.$$

We have used that the real diagonal matrix A_n commutes with L. Consequently

$$\lim_{n \to \infty} f_n(z) = \lim A_n^{-1} f A_n(z) = L(z) \quad \text{uniformly on compact subsets of } \mathbb{C}.$$

In short, L is nothing but the "blow-up" of f at $z = 0$.

It now follows that

$$\lim A_n^{-1} G A_n \ = \ \lim A_n^{-1} g_n G g_n^{-1} A_n \ = \ B^{-1} G B.$$

This implies that the sequence of groups $\{A_n^{-1}GA_n\}$ *converges geometrically*. In the next chapter, we will study this notion in detail; suffice it to say that every $B^{-1}gB$ is the limit of elements of the approximants $\{A_n^{-1}GA_n\}$, namely $B^{-1}gB = \lim(A_n^{-1}g_n)g(g_n^{-1}A_n)$. Conversely, the limit of any convergent sequence of elements of $\{A_n^{-1}GA_n\}$ lies in $B^{-1}GB$, namely,

$$h := \lim A_n^{-1}h_nA_n = \lim A_n^{-1}g_n(g_n^{-1}h_ng_n)g_n^{-1}A_n = B^{-1}(\lim g_n^{-1}h_ng_n)B.$$

Recall that for any $g \in G$, $f \circ g = \phi(g) \circ f$. Given $g \in G$,

$$\begin{aligned} L \circ B^{-1}gB \circ L^{-1} &= \lim_{n\to\infty} A_n^{-1}f \circ g_ngg_n^{-1} \circ f^{-1}A_n \\ &= \lim_{n\to\infty} A_n^{-1}\phi(g_ngg_n^{-1}) \circ f \circ f^{-1}A_n = \lim_{n\to\infty} A_n^{-1}\phi(g_ngg_n^{-1})A_n. \end{aligned}$$

The element on the left is therefore a Möbius transformation. We have established that $L \circ h \circ L^{-1}$ is a Möbius transformation for any $h \in B^{-1}GB$.

Since not all elements of G fix $B(\infty)$, there exists $h \in B^{-1}GB$ with $h(\infty) \neq \infty$. We now know that LhL^{-1} is a Möbius transformation. We claim that this forces L itself to be a Möbius transformation, necessarily fixing 0 and ∞.

For let ℓ be a euclidean line in $\mathbb{C}$. Then $L^{-1}(\ell)$ is again a straight line. Choose ℓ such that $L^{-1}(\ell)$ does not go through $h^{-1}(\infty)$. Then $h \circ L^{-1}(\ell) = C$ is a proper circle. Therefore $LhL^{-1}(\ell) = L(C)$ is a circle as well, since on the one hand L maps bounded sets to bounded sets, and on the other, LhL^{-1} is Möbius. But an affine mapping L that by definition fixes 0 and ∞ cannot send a circle onto a circle unless it can be expressed as $z \mapsto az$ (or $z \mapsto a\bar{z}$, if we allowed f and hence L to be orientation reversing). So L is a Möbius transformation, as claimed, and it has the simple form $z \mapsto az$.

Pick three distinct points $p_1, p_2, p_3 \in \mathbb{S}^2$. For any homeomorphism $F : \mathbb{S}^2 \to \mathbb{S}^2$ set $N(F) = F^\sharp \circ F$ where the Möbius transformation $F^\sharp$ is uniquely chosen so that $N(F)$ fixes each p_i. Upon setting $u_n = g_n^{-1}A_n$ so that $\lim u_n = B$,

$$N(f_n) = N(A_n^{-1}fA_n) = N(fA_n) = N(fg_nu_n) = N(\phi(g_n)fu_n) = N(fu_n).$$

Going to the limit,

$$N(L) = \lim_{n\to\infty} N(f_n) = N(fB).$$

Since L and B are Möbius transformations, f must be one as well. □

Mostow's Rigidity Theorem follows from this result. We have to construct, given the isomorphism $\phi : G \to H$, a quasiconformal automorphism of $\mathbb{S}^2$ that induces it. If $\mathcal{M}(G)$ has finite volume, results of Waldhausen [1968, Lemma 6.3, Theorem 6.1] applied and extended to the noncompact case in [Marden 1974a], or of Tukia [1985b, Theorem 4.7] show that there is an orientation preserving or reversing quasiconformal mapping Φ_* between the manifolds, inducing ϕ on the fundamental group, so that $\mathcal{M}(H)$ has finite volume as well. By replacing H by JHJ if necessary we may assume it is an ordinary quasiconformal mapping. A lift Φ to $\mathbb{H}^3$ is a quasiconformal mapping, and quasiconformal mappings of, say, upper half-space extend to be quasiconformal on $\mathbb{C} \cup \infty$ [Gehring 1962]. Also $\Phi G \Phi^{-1} = H$. Quasiconformal maps of $\mathbb{S}^2$ are differentiable with nonzero derivative almost everywhere [Ahlfors 1966]. Since G is geometrically finite and $\partial \mathcal{M}(G) = \varnothing$, every point on $\mathbb{S}^2$ is a limit point, and all those except the countable number of parabolic fixed points are conical limit points. All that remains is to apply Tukia's result at one point of differentiability which is not a parabolic fixed point.

If $\mathcal{M}(G)$ is a closed manifold the following alternate argument can be used: It follows from topology that there is a *homotopy equivalence* (see the discussion in Section 5.1) between the manifolds: continuous maps $f_1 : \mathcal{M}(G) \to \mathcal{M}(H)$ and $f_2 : \mathcal{M}(H) \to \mathcal{M}(G)$ such that $f_1 \circ f_2$ and $f_2 \circ f_1$ are homotopic to the identity. Working in terms of a piecewise linear structure (subdividing into hyperbolic tetrahedra) on the manifolds, the mappings can be taken to be Lipschitz. It turns out [Mostow 1973, Lemma 9.2; Thurston 1979, p. 5.39] that their lifts F_1, F_2 to $\mathbb{H}^3$ are quasiisometries; see Exercise 3-19. Consequently each one extends to $\partial \mathbb{H}^3$ and is quasiconformal there. Using the cusp tori, this approach too extends to the finite volume case; see [Prasad 1973; Thurston 1979, p. 5.39; Tukia 1985b, Lemma 3.4]. □

It is interesting to compare the situation we just considered to the case of a quasiconformal map $f : \mathbb{H}^2 \to \mathbb{H}^2$. Likewise f can be extended to $\mathbb{S}^1$ and is again a homeomorphism there, necessarily having a derivative almost everywhere. The extension to $\partial \mathbb{H}^2$ is either the restriction of a Möbius transformation, or its derivative is zero wherever it exists. Now one knows in advance that most fuchsian groups have nontrivial deformations. The corresponding homeomorphisms of the circle $\partial \mathbb{H}^2$ are therefore examples of totally singular functions: their derivatives are zero almost everywhere. This seems to be the simplest construction of singular functions.

3.14 Exercises and explorations

3-1. (a) Prove the area formula for a surface S of constant gaussian curvature $K = 0, \pm 1$, the area being given in the euclidean, spherical, or hyperbolic metric. Here S is a closed surface of genus $g \geq 0$, with $n \geq 0$ punctures and $m \geq 0$ cone points of orders $\{2 \leq r_i < \infty\}$.

$$K \, \mathrm{Area}(S) = 2\pi \left(2 - 2g - n - \sum_{i=1}^{m} \left(1 - \frac{1}{r_i} \right) \right). \tag{3.9}$$

Hint: The Euler characteristic formula for a triangulated closed surface is $\chi(S) = T - E + V = 2 - 2g$, where T is the number of triangles, E the number of edges, V the number of vertices and g the genus. Thus if S is a closed surface of genus $g \geq 0$, the Euler characteristic is $\chi(S) = 2 - 2g$, while if S is a closed surface of genus g with n punctures, $\chi(S) = 2 - 2g - n$ (the punctures are not counted as vertices).

Now we will compute the area. Cut the surface into small geodesic triangles. Each puncture and cone point should be a vertex. Think of how the neighborhood on each arises by projection from the branched universal cover. Since each triangle has three edges each of which is shared by the adjacent triangle, $2E = 3T$. The area of each triangle satisfies $K\,\mathrm{Area}(\Delta) = \theta_1 + \theta_2 + \theta_3 - \pi$. If there are no punctures or cone points, summing the triangles we find that $K\,\mathrm{Area}(S) = 2\pi V - \pi T = 2\pi\chi(S)$. If there are cusps (cone points of order ∞ on a negatively curved surface) the area is too great because the angle sum about a cusp is 0 instead of 2π, so $2\pi n$ must be subtracted. At a cone point the angle sum is instead $2\pi/r_i$ rather than 2π, so we must subtract the difference $2\pi(1 - 1/r_i)$.

Thus if S is a closed surface of genus g with n-punctures, and $K = -1$, we have

$$\mathrm{Area}(S) \geq 2\pi|\chi(S)|. \tag{3.10}$$

(b) Conversely, prove that every possibility S allowed by (3.9) can be realized as $S = P/G$ where P is exactly one of $\mathbb{S}^2$, $\mathbb{C}$, $\mathbb{H}^2$ and G is a group of Möbius transformations acting on P (if there are no cone points, this is a consequence of the Uniformization Theorem, Section 2.6).

Hint: To simplify notation, consider the case of $\mathbb{S}^2$ with n cone points $\{\zeta_i\}$ of corresponding orders $\{r_i\}$. Fix a point O and take n simple loops $\{\gamma_i\}$ from O, each surrounding exactly one cone point, and mutually disjoint except at O. Let H be the normal subgroup of the fundamental group of $S = \mathbb{S}^2 \setminus$ cone points generated by the loops $\{\gamma_i^{r_i}\}$. Let $\tilde{S}$ denote the normal covering Riemann surface determined by H. Each lift of each $\gamma_i^{r_i}$ is a simple loop retractable to a puncture. The group of cover transformations is isomorphic to $\pi_1(S)/H$. In particular each lift of γ_i determines an element of order r_i that necessarily extends to and fixes the corresponding puncture. There are a countable number of such lifts. When the punctures are added to $\tilde{S}$ we obtain a simply connected Riemann surface $\tilde{S}^*$. Now apply the Uniformization Theorem to $\tilde{S}^*$. In general $\tilde{S}^*$ is noncompact and conformally equivalent to $\mathbb{H}^2$: there are only a finite number of configurations that lead to the plane or sphere.

(c) For a closed, oriented surface S of genus g with riemannian metric h and Gaussian curvature $K(h)$ the Gauss–Bonnet formula reads

$$2\pi\chi(S) = 2\pi(2-2g) = \iint_S K(h)\,dA_h,$$

where $\chi(S)$ is the Euler characteristic and dA_h is the element of surface area. In the hyperbolic case $K(h) = -1$ and the surface area is $4\pi(g-1)$, $g \geq 2$. Hyperbolic metrics are best: If $K(h) \geq -1$ (resp. ≤ -1), then $\mathrm{Area}_h(S) \geq \mathrm{Area}_{\mathrm{hyp}}$ (resp. $\leq$), with equality only when h is the hyperbolic metric. For generalizations to 3-manifolds, see [Besson et al. 1999; Storm 2002a; 2002b].

Equation (3.9) is sometimes applied to cone manifolds with arbitrary cone angles. It holds for cone angles $2\pi/r_i \leq 2\pi$. If all the cone points on an n-punctured surface S of genus g have angles instead satisfying $2\pi/r_i \geq 2\pi$, Equation (3.10) becomes

$$\text{Area}(S) \leq 2\pi\,|\chi(S)|.$$

A common application of the area formula is to find the possibilities that a closed surface of genus g and n punctures with designated cone points carries the spherical metric $K = +1$, euclidean metric ($K = 0$), or hyperbolic ($K = -1$). In other words, the surface is covered by $\mathbb{S}^2$, $\mathbb{C}$ or $\mathbb{H}^2$. This gives rise to three inequalities:

$$2g + n + \sum_{i=1}^{m}\left(1 - \frac{1}{r_i}\right) < 2 \qquad \text{spherical case,} \tag{3.11}$$

$$2g + n + \sum_{i=1}^{m}\left(1 - \frac{1}{r_i}\right) = 2 \qquad \text{euclidean case,} \tag{3.12}$$

$$2g + n + \sum_{i=1}^{m}\left(1 - \frac{1}{r_i}\right) > 2 \qquad \text{hyperbolic case.} \tag{3.13}$$

Inequality (3.11) requires $g = 0$ and $n = 0, 1$. If $n = 1$ then $m = 1$ and $2 \leq r_1 < \infty$. For $n = 0$, if $m = 3$ the possibilities for the cone points are $(2, 3, 5)$, $(2, 3, 4)$, $(2, 3, 3)$, $(2, 2, n)$; if $m = 2$ then $2 \leq r_1, r_2 < \infty$; if $m = 1$, then $2 \leq r_1 < \infty$.
Equality (3.12) requires $g = 0, 1$. If $g = 1$ then $m, n = 0$. For $g = 0$, we can have $n = 2$ and $m = 0$; otherwise if $n = 1$, then $m = 2$ and $r_1 = r_2 = 2$; if $n = 0$, the cone points are given by $(2, 2, 2, 2)$, $(3, 3, 3)$, $(2, 3, 6)$, or $(2, 4, 4)$.
Inequality (3.13) is satisfied by all combinations except those listed already.

(d) Show that as $\Gamma \subset \text{PSL}(2, \mathbb{R})$ ranges over all fuchsian groups (that may have elliptics and/or parabolics), the area of $\mathbb{H}^2/\Gamma$ achieves its minimum value $\frac{\pi}{21}$ uniquely for the $(2, 3, 7)$-triangle group. Conclude that for a group R representing a closed surface R of genus $g \geq 2$, the order of the group $C(R)$ of conformal automorphisms of R cannot exceed $4\pi(g-1)/(\pi/21) = 84(g-1)$. (Note that $C(R)$ is isomorphic to $N(G)/G$, where $N(G)$ is the normalizer of G, and the area of $\mathbb{H}^2/N(G)$ is not less than $\pi/21$.)
Show that if the fuchsian group G is of finite index n in the fuchsian group H, the area of $\mathbb{H}^2/H$ is n times that of $(\mathbb{H}^2/G)$. Because n cannot become too large, conclude that every fuchsian group of finite area is contained in a *maximal fuchsian group*, one that has finite area and is not a subgroup of any other fuchsian group.
Does the same argument work for finite volume kleinian groups (Section 4.11.1)?

3-2. If A is loxodromic prove that in the hyperbolic metric $\min_{\vec{x} \in \mathbb{H}^3} d(\vec{x}, A(\vec{x}))$ is achieved only when $\vec{x}$ lies on the axis of A.

Exercise 1-4 showed that the set $V = \{\vec{x} \in \mathbb{H}^3 : d(\vec{x}, A(\vec{x})) < \epsilon\}$, if nonempty, is a radius ϵ tube about the axis of A.

Show that if A is parabolic, the set $\{\vec{x} \in \mathbb{H}^3 : d(\vec{x}, A(\vec{x})) < \epsilon\}$ is a horoball at the fixed point of A.

3-3. Prove that for a fuchsian group G, the universal horodisk at a parabolic fixed point is not penetrated by the axis of any loxodromic element that represents a simple geodesic on $\mathbb{H}^2/G$. Is the same statement true for the universal horoball in a kleinian group? (Hint: apply $z \mapsto z+1$ to the axis).

Prove that a discrete group with all real traces is conjugate to a fuchsian group.

3-4. Show that if $x \in \mathbb{H}^3$ approaches $z \in \partial\mathbb{H}^3$, then the limit of the half-space H_g is the half-space determined as follows. There is a unique horosphere σ at z such that σ is tangent to the horosphere $g^{-1}\sigma$ at $g^{-1}(z)$. Take the hyperbolic plane tangent to both horospheres at their point of tangency; it is orthogonal to the geodesic with endpoints z, $g^{-1}(z)$. Choose the half-space determined by this plane that is adjacent to z. If $z = \infty$, this half-space is the exterior of the isometric hemisphere for g. Also see Lemma 1.5.4.

3-5. *Figure*-8 *knot.* Find a Dirichlet region for the rank-two parabolic group

$$G = \langle z \mapsto z+1,\ z \mapsto z+\tau;\ \operatorname{Im}\tau > 0\rangle.$$

Show that it has generically six edges, but in some situations it has only four. The square and the regular hexagon provide the associated torus with symmetries of order four and order six.

Compute the hyperbolic volume of the part of the polyhedron lying above a horosphere (a horizontal plane, in the present situation). Show that the quotient $\mathbb{H}^3 \cup \mathbb{C}/G$ is homeomorphic to $\{0 < |z| \leq 1,\ z \in \mathbb{C}\} \times \mathbb{S}^1$, that is, the complement of the central circle in the solid torus. This is the prototype of the local structure about a knot when the knot complement has a hyperbolic structure — as all of them do, except torus knots and satellite knots (Section 6.3). The parabolic fixed point is "stretched" into the knot.

For example, the figure-8 knot can be formed as follows. In the upper half-space model, choose an ideal tetrahedron with one vertex at ∞, as in Exercise 1-22. Each of the four faces is an ideal triangle. The ideal vertices of each face lie on a circle in $\mathbb{S}^2$.

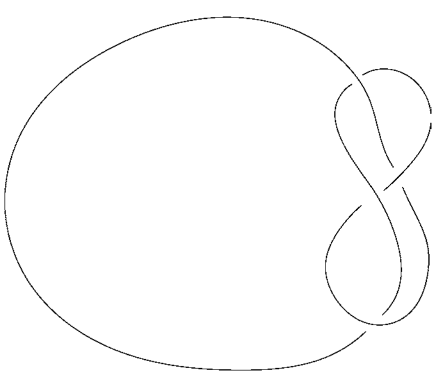


Fig. 3.8. The figure-8 knot.

The circles corresponding to adjacent faces intersect, and their angle of intersection is the dihedral angle between the faces. Arrange it so that the six dihedral angles are all 60° so as to become the regular ideal tetrahedron (compare Exercise 1-23). In fact the dihedral angles of any ideal tetrahedron add up to 360°.

Line up two such ideal tetrahedra T_1 and T_2, one next to the other so they share a face and the ideal vertex ∞. There are six free faces on the union of the two tetrahedra. The faces can be paired and the face identification via isometries precisely given so that the tetrahedral union is a fundamental polyhedron for the group G generated by the face pairing transformations and $\mathbb{H}^3/G$ is homeomorphic to the complement of the figure-8 knot in $\mathbb{S}^3$. The five ideal vertices become parabolic fixed points which are in the single parabolic conjugacy class of G. For details see [Thurston 1997, pp. 39–42], [Ratcliffe 1994, §10.5], or [Neumann 1999].

The figure-8 knot complement is $\mathbb{H}^3/G$, where G can be taken to be generated by

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & e^{\pi i/3} \\ 0 & 1 \end{pmatrix}.$$

For a discussion of hyperbolic knots see Section 6.3.

3-6. *Volume of maximal solid cusp tori* [Adams 1987]. We will use the fact that the densest circle packing in the plane with circles of the same radius is the hexagonal packing: each circle is surrounded by six others. This is applied as follows.

Suppose P is a parallelogram. Place a disk of radius r centered at each of the four vertices. Assume that the interiors of the disks are mutually disjoint. If $|P|$ denotes the area of P, then $|P| \geq 2r^2\sqrt{3}$. Equality occurs if and only if all the sides of P have the length $2r$. If the sides of P have length ≥ 1, then $r \geq 1/2$. If one side has length one, $r = 1/2$.

Now consider a hyperbolic 3-manifold $\mathcal{M}(G)$ such that G has a rank two parabolic subgroup. We can conjugate so that $G_\infty = \langle z \mapsto z+1, z \mapsto z+\tau \rangle$ with $\tau = u+iv$, $-\frac{1}{2} \leq u \leq \frac{1}{2}$, $y \geq \sqrt{3}/2$ (Exercise 2-5). A *maximal horoball* at the fixed point $z = \infty$ is the largest horoball $\mathcal{H}_\infty$ with the property that $g(\mathcal{H}_\infty) \cap \mathcal{H}_\infty = \varnothing$ for all $g \in G$, $g \notin G_\infty$. In our setup with ∞ the fixed point, this means that $\mathcal{H}_\infty = \{(z,t) \in \mathbb{H}^3 : t > s\}$ with the smallest possible $1 \geq s > 0$.

Let σ denote the horosphere $\{(z,t) : t = s\}$. Because this bounds the maximal horoball at ∞ there will be an element $g \in G$, $g \notin G_\infty$ such that the euclidean sphere $g(\sigma)$ is tangent to σ. We may assume that $g(\sigma)$ is based at $z = 0 = g(\infty)$.

Prove that $g^{-1}(\infty)$ cannot lie in the orbit $G_\infty(0)$. Hint: suppose otherwise so that for some $h \in G_\infty$, $h^{-1}g^{-1}(\infty) = 0$. Also $h^{-1}g^{-1}(0) = \infty$. Therefore $h^{-1}g^{-1}$ fixes a point on the vertical half line, a contradiction since G has no elliptic elements.

In $\mathbb{C}$, choose the fundamental parallelogram P for G_∞ to have vertices at 0, 1, τ, and $1+\tau$. The horoball $g(\mathcal{H}_\infty)$ is tangent to $\mathbb{C}$ at $z = 0$ and its G_∞-orbit contains horoballs tangent to $\mathbb{C}$ at all the vertices of P. Its G_∞-orbit is also disjoint from the G_∞-orbit of the horoball $g^{-1}(\mathcal{H}_\infty)$. There will be at least one point $\zeta \in P$ which is a tangent point of the latter G_∞-orbit.

All these horoballs tangent to $\mathbb{C}$ have the same euclidean radius $r = 1/(2s)$. The vertices of P have distance at least $2r = 1$ apart, and also must be distance at least 1 from ζ. Place disks of radius $1/2$ centered at the vertices of P and at ζ. Consider their G_∞-orbit. Their interiors are mutually disjoint and P has to be covered by the equivalent of two disks. Deduce that $|P| \geq \sqrt{3}$.

The volume of $\mathcal{H}_\infty/G_\infty$ is $|P|/2s^2$.

Conclude that *the volume of the maximal solid cusp torus is at least* $\sqrt{3}/2$ *in any hyperbolic manifold* $\mathcal{M}(G)$ *with nonelementary* G.

Compare with Exercise 2-10.

Now show that the number of primitive lattice points of the orbit $G_\infty(0)$ whose distance from $z = 0$ is less than 2π is ≤ 48 [Bleiler and Hodgson 1996]; there are at most 24 simple closed geodesics in the quotient torus of length $< 2\pi$. Primitive means that the ray from 0 to the lattice point does not pass through any other lattice points.

3-7. *Subgroups of geometrically finite groups* [Thurston 1986b]. Suppose $\mathcal{M}(G)$ is a geometrically finite hyperbolic 3-manifold such that its convex hull $\mathcal{C}(G) \neq \mathcal{M}(G)$. Prove that for every finitely generated subgroup G_1 of G, $\mathcal{M}(G_1)$ is also geometrically finite.

Hint: Consider first the case that G has no parabolics. Then $\mathcal{C}(G)$ is compact. There exists d such that every point $x \in \mathcal{C}(G)$ has distance at most d from $\partial\mathcal{C}(G)$. Let $\widehat{\mathcal{C}}(G)$ denote the lift to $\mathbb{H}^3$. Every point in $\widehat{\mathcal{C}}(G)$ has distance at most d from $\partial\widehat{\mathcal{C}}(G)$. The quotient $\widehat{\mathcal{C}}(G)/G_1$ is a convex submanifold of $\mathcal{M}(G_1)$ and hence it contains $\mathcal{C}(G_1)$. Each point $x \in \mathcal{C}(G_1)$ therefore has distance at most d from $\partial\widehat{\mathcal{C}}(G)/G_1$ and then also from $\partial\mathcal{C}(G_1)$. Now the Ahlfors Finiteness Theorem implies that $\partial\mathcal{C}(G_1)$ has a finite number of components and each component is a compact surface without boundary (there are no parabolics in G_1). Consequently $\mathcal{C}(G_1)$, being covered by a finite number of d-balls with centers on $\partial\mathcal{C}(G_1)$, is compact. Therefore G_1 is geometrically finite.

The proof in the general case also uses the thick/thin decomposition of $\mathcal{C}(G)$.

3-8. *Klein–Maskit combination theory.* Here we will display only the classical situations. In [Maskit 1988] the reader will find extensive generalizations.

(i) Suppose G is a kleinian group. Select two mutually disjoint closed disks D_1, D_2 in $\Omega(G)$ such $g(D_i) \cap D_j = \varnothing$ for $i, j = 1, 2$ and for all $g \neq \mathrm{id} \in G$. Let T be any Möbius transformation that maps the exterior of D_1 onto the interior of D_2. Prove that $G^* = \langle G, T\rangle$ is also discrete, as claimed in Section 3.7.

Topologically show that what you have done is the following. The projection $\pi : D_i \to \Delta_i$ to the quotient $\mathcal{M}(G)$ is a homeomorphism. Remove Δ_1, Δ_2 from $\partial\mathcal{M}(G)$ and identify the resulting boundaries $\partial\Delta_1$ and $\partial\Delta_2$. If the two disks lie on the same component of $\partial\mathcal{M}(G)$ what you have done is create a new handle. If they lie in different boundary components, you have connected the two components. Otherwise

$\partial\mathcal{M}(G^*)$ is the same as $\partial\mathcal{M}(G)$. In either case, the simple loop $\partial\Delta_1 = \partial\Delta_2$ on $\partial\mathcal{M}(G^*)$ bounds a disk within $\mathcal{M}(G^*)$. This disk *does not* divide the 3-manifold.

Exactly the same process can be used to connect two manifolds $\mathcal{M}(G_1)$ and $\mathcal{M}(G_2)$. In this case $G^* = G_1 * G_2$ is a free product since the new disk divides.

Show that you can adjoin a solid torus and/or a solid cusp torus to $\mathcal{M}(G)$.

What happens if you make the following alternative combination? Given a small closed disk $D \subset \Omega(G)$, let J denote reflection in the circle ∂D, and equally in the plane in $\mathbb{H}^3$ rising from ∂D. Form the new group $G^* = \langle G, JGJ\rangle$. This too will be discrete. Describe $\mathcal{M}(G^*)$.

(ii) Suppose ζ_1, ζ_2 are two parabolic fixed points of G. Suppose $\Delta_1 \in \Omega(G)$ is a closed horodisk associated with ζ_1 and Δ_2 is a closed horodisk associated with ζ_2 so that $g(\Delta_i) \cap \Delta_i = \varnothing$, $i = 1, 2$, unless $g \in \mathrm{Stab}_{\zeta_i}$. Possibly $\zeta_1 = \zeta_2$ and then the disks are externally tangent at the fixed point. Let B be any Möbius transformation that maps the exterior of Δ_1 onto the interior of Δ_2 *and* conjugates $\mathrm{Stab}(\zeta_1)$ to $\mathrm{Stab}(\zeta_2)$. Prove that $G^* = \langle G, B\rangle$ is discrete. Topologically what has happened is this: We have chosen two circles c_1, c_2 about two distinct punctures on $\partial\mathcal{M}(G)$, the projections of the two horocycles. Remove the once punctured disks bounded by these two circles from $\partial\mathcal{M}(G)$ and identify the two circles. If the two circles are on the same boundary component of $\mathcal{M}(G)$, that component loses two punctures but gains a handle. If they are on different components, the two components become connected and lose a puncture each. In G^* the two cyclic parabolic groups become conjugate; if $\zeta_1 = \zeta_2$ the cyclic parabolic group becomes a rank two parabolic group. See also Exercise 4-18.

Algebraically G^* is the free product with amalgamation of the two cyclic parabolic groups. Likewise the construction can be carried out to join two different manifolds.

It was originally hoped that with the two classical combination techniques, (i) and (ii), all kleinian groups could be constructed. Peter Scott, in the mid 1970s, showed (personal communication) that this cannot be the case. It is a key part of the Thurston theory, specifically the skinning lemma (Section 6.2), that allows general forms of combination to be effectively and generally applied — it shows that the group can be deformed so that there exist Möbius transformations that do the job required of hyperbolic gluing. Armed with the skinning lemma, most kleinian groups can be formed from simpler ones using hyperbolic gluing — the combination theorems.

3-9. *Extended quasifuchsian groups.* Let G be a fuchsian or quasifuchsian group with $\Omega(G) = \Omega_1 \cup \Omega_2$ the components of the regular set. Suppose there is a Möbius transformation T that maps Ω_1 onto Ω_2 and such that if $g \in G$ then also $TgT^{-1} \in G$. Show that the extended group $G^* = \langle G, T\rangle$ is discrete. Describe the topology of $\mathcal{M}(G^*)$ (it has only one boundary component).

The group G^* is called an *extended fuchsian* or *extended quasifuchsian* group. It has the same limit set as G and an index two subgroup which is fuchsian or quasifuchsian. Construct an extended fuchsian group by adjoining $z \mapsto -z$ to the modular group.

3-10. Suppose G is a finitely generated kleinian group and Ω is a simply connected component of $\Omega(G)$ with the following properties:

(i) Ω is invariant under G and is a proper subset of $\Omega(G)$.
(ii) Every simple loop in $S = \Omega/G$ that determines a parabolic element of G is retractable in S to a puncture.

Apply the Ahlfors Finiteness Theorem and the Cylinder Theorem to prove that G is a fuchsian or quasifuchsian group.

3-11. *Function and Schottky groups; compression bodies.* Assume that a component Ω of $\Omega(G)$ is invariant under G. Traditionally, complex analysts have called such a group a *function group* because by using Poincaré series, differentials and functions can be constructed on it. We will however reserve the name to the cases that Ω is not simply connected. Here we will assume that G is finitely generated without elliptics.

An orientable, compact, irreducible 3-manifold M^3 is called a *compression body* by the topologists if it has a boundary component $S \subset \partial M^3$ for which the inclusion $\pi_1(S) \hookrightarrow \pi_1(M^3)$ is surjective. It is referred to as a *trivial* compression body if $M^3 = S \times [0, 1]$; we will not use the term for this case.

If $\mathcal{M}(G)$ is compact, it can be described topologically as the result of taking a 3-ball B, cutting $n \geq 2$ holes in ∂B and attaching to the boundary of the holes, the boundary curves of the following collection: solid tori, each with one hole cut out the boundary, and closed surface bundles $S_k \times [0, 1]$, with the genus of S_k exceeding one, each with a hole taken out of a boundary component. In the opposite direction, on the compressible boundary component S there is a finite system of nontrivial simple loops that bound disks in $\mathcal{M}(G)$. Cut $\mathcal{M}(G)$ along these disks to get one or more pieces M_i. Here we are using Dehn's Lemma and the Loop Theorem. If there is only one piece, then it is a ball and $\mathcal{M}(G)$ is a handlebody. Otherwise each $M_i \cong S_i \times [0, 1]$ where S_i is an incompressible boundary component of $\mathcal{M}(G)$.

When there are parabolics and G is geometrically finite, compactify $\mathcal{M}(G)$ by removing solid cusp tori and solid pairing tubes. Then the analysis is the same. Note that in this case some of the components may be essentially solid cusp tori themselves. Algebraically, G is the free product of closed surface groups and cyclic groups. If G is an N-generator function group, find estimates for the number of pieces, and the genus and punctures of each [Marden 1974a].

Prove that a general geometrically finite manifold $\mathcal{M}(G)$ can be decomposed along incompressible surfaces into finitely many compression bodies and submanifolds with incompressible boundary. This is a result of Bonahon [1986], who also showed that the decomposition is unique up to isotopy. Two compact compression bodies with isomorphic fundamental groups are homeomorphic [McCullough and Miller 1986].

The simplest example is a (not necessarily classical) Schottky group representing a handlebody $\mathcal{M}(G)$ of genus $g \geq 1$. Equally important for a Schottky group are the simple loops which are not compressing. Show that there exist simple noncompressing loops that divide the surface into two parts; see Exercise 5-16.

In [McCullough and Miller 1986] it is proved after a long argument that given a compressible boundary component S of a geometrically finite $\mathcal{M}(G)$ without parabolics, there is a submanifold $X \subset \mathcal{M}(G)$ with the following properties: (i) S is a boundary component of X, (ii) $\partial X \setminus S$ is incompressible in $\mathcal{M}(G)$, and (iii) the image of the inclusion $\pi_1(S) \hookrightarrow \pi_1(\mathcal{M}(G))$ is precisely $\pi_1(X)$, that is, X is a compression body with S its compressible boundary component.

Suppose $\mathcal{M}(G)$ is compact with an incompressible boundary component S. Show that $\pi_1(S)$ either has index at most two in G (Exercise 3-9), or it has infinite index in G [Hempel 1976, Theorem 10.5].

3-12. Assume that $\infty \in \Omega(G)$ and Ω is a component of $\Omega(G)$ with $\infty \neq \Omega$. Suppose there exists a *relatively compact* fundamental set F for the action of $\mathrm{Stab}(\Omega)$ (recall that this is the group $\{g \in G : g(\Omega) = \Omega\}$). Prove that

$$\mathrm{Diam}(\Omega)^2 \leq \sum_{g \in \mathrm{Stab}(\Omega)} \mathrm{Diam}(g(F))^2.$$

Prove further that if $\{g_i(\mathrm{Stab}(\Omega))\}$ is the set of left cosets of $\mathrm{Stab}(\Omega)$ in G, then

$$\sum \mathrm{Diam}(g_i(\Omega))^2 < \infty.$$

Here Diam is the euclidean diameter of the set. There is a one-to-one correspondence between left or right cosets of $\mathrm{Stab}(\Omega)$ in G and components of the orbit $G(\Omega)$.

3-13. *Boundary fixed points* [Maskit 1974]. Suppose H is such that the quotient $\Omega(H)/H$ has a finite number of components each of which is a closed surface. Prove that if ∞ is not a limit point,

$$\sum_{h \in H} |c_h|^4 < \infty,$$

where $|c_h|^{-1}$ is the radius of the isometric circle of $h \neq \mathrm{id}$. *Hint:* the orbit of a fundamental region has finite spherical area since there is no overlap.

Now suppose G is a kleinian group, ∞ is not a fixed point, and $\Omega \subset \Omega(G)$ is a component of the regular set. Consider $\mathrm{Stab}(\Omega) = \{g \in G : g(\Omega) = \Omega\}$. Assume that the quotient $\Omega/\mathrm{Stab}(\Omega)$ is a closed surface. Let $\{\Omega_i\}$ denote the components of the G-orbit of Ω; that is, if $G = \bigcup g_i \,\mathrm{Stab}(\Omega)$ is the coset decomposition, then we can take $\Omega_i = g_i(\Omega)$. Prove for the spherical diameters that $\sum \mathrm{Diam}^4(\Omega_i) < \infty$. *Hint:* $\mathrm{Diam}(g_i(\Omega_i)) \leq |c_i|^{-2} d_i^{-1}$ where d_i is the spherical distance between $g_i^{-1}(\infty)$ and Ω.

Deduce that if a loxodromic $g \in G$ has a fixed point on $\partial\Omega$ then $g^k(\Omega) = \Omega$ for some k. *Hint:* If no power g^k preserves Ω then $\sum \mathrm{Diam}(g^k(\Omega)) = \infty$.

The same conclusion holds in the more general case that $\Omega/\mathrm{Stab}(\Omega)$ has in addition a finite number of punctures.

More generally, prove the following result from [Anderson 1994]. Suppose G is a not necessarily finitely generated group but $G_1 \subset G$ is a finitely generated subgroup. Assume the loxodromic $\gamma \in G$ has a fixed point in $\Lambda(G_1)$. Prove that $\gamma^k \in G_1$ for some $k \geq 1$.

Analyze the following case. Suppose Ω is a component of the regular set of the nonelementary, torsion-free finitely generated kleinian group G. Assume the loxodromic $A \in \mathrm{Stab}(\Omega)$ represents a simple closed curve c on $\Omega/\mathrm{Stab}(\Omega)$. Suppose $A = B^n$, $n \geq 2$, where $B \in G$ preserves $\Omega_1 \neq \Omega$ and also represents a simple closed curve c'. Suppose c and c' are disjoint in the quotient. Now B^n preserves both Ω and Ω_1. In $\mathcal{M}(G)$, c is freely homotopic to c'^n. Since c and c' are disjoint curves, by the cylinder theorem there is a simple curve c^* near c' that is freely homotopic to c. When can you conclude that $n = \pm 1$?

3-14. *Commensurability.* Two subgroups groups Γ_1, Γ_2 of a larger group G^* (which will usually be $\mathrm{PSL}(2,\mathbb{C})$ or $\mathrm{PSL}(2,\mathbb{R})$) are said to be *commensurable* ($\Gamma_2 \sim \Gamma_1$) if the subgroup of common elements $\Gamma_2 \cap \Gamma_1$ is of finite index in both Γ_1 and Γ_2. Prove that if Γ_1 is geometrically finite, Γ_2 is as well [Greenberg 1977].

In the context of kleinian groups, $\Gamma_2 \sim \Gamma_1$ if and only if $\mathcal{M}(\Gamma_1 \cap \Gamma_2)$ is a finite-sheeted cover of both $\mathcal{M}(\Gamma_1)$ and $\mathcal{M}(\Gamma_2)$.

The *commensurability group* or *commensurator* $C(\Gamma)$ of a kleinian group Γ is the group $C(\Gamma) = \{g \in \mathrm{PSL}(2,\mathbb{C}) : g\Gamma g^{-1} \sim \Gamma\}$.

Prove the following special case of [Greenberg 1974, Theorem 2(4)]. If Γ is a finitely generated, nonelementary group whose limit set $\Lambda(\Gamma)$ is not a round circle on $\mathbb{S}^2$ nor is all $\mathbb{S}^2$, then the index $[C(\Gamma) : \Gamma]$ is finite.

To establish this, show that $C(\Gamma)$ is discrete. Then show it has the same limit set as Γ. Therefore if F is a fundamental region for $C(\Gamma)$, and $\{g_i\}$ is a set of coset representatives for Γ in $C(\Gamma)$, show that $\bigcup g_i(F)$ is a fundamental set for Γ on $\Omega(\Gamma)$. Applying the Ahlfors Finiteness Theorem, $\{g_i\}$ is a finite set and therefore $[C(\Gamma) : \Gamma] < \infty$. Alternatively use the fact that the group K of elements that map $\Lambda(G)$ onto itself is discrete: As a closed subgroup of $\mathrm{PSL}(2,\mathbb{C})$, the identity component of K is a connected Lie subgroup. Since it is not $\mathrm{PSL}(2,\mathbb{C})$ or conjugate to $\mathrm{PSL}(2,\mathbb{R})$ it either has a common fixed point in $\mathbb{H}^3 \cup \partial\mathbb{H}^3$, or it is the identity [Greenberg 1977], implying that K is discrete.

Prove that $C(\Gamma)$ contains any group H with the same limit set as Γ. That is $C(\Gamma)$ is the group of all Möbius transformations that map the ordinary set $\Omega(G)$ onto itself (or equivalently, map $\Lambda(G)$ onto itself). Therefore, if H contains Γ, Γ has finite index in H.

If $\mathcal{M}(G_1)$, $\mathcal{M}(G_2)$ have finite volume, show that G_1, G_2 are commensurable if and only if there are *isomorphic* subgroups of finite index $H_1 \subset G_1$ and $H_2 \subset G_2$.

What do your results say about the normalizer $N(\Gamma)$ of Γ in $\mathrm{PSL}(2,\mathbb{C})$?

In contrast to our analysis, if $\Gamma = \mathrm{PSL}(2,\mathbb{Z})$ and $G^* = \mathrm{PSL}(2,\mathbb{R})$, then $C(\Gamma)$ contains $\mathrm{PSL}(2,\mathbb{Q})$ so that $[C(\Gamma) : \Gamma] = \infty$ and $C(\Gamma)$ is dense in G^*.

3-15. *Finiteness theorems.* Suppose that G is a finitely generated kleinian group. Prove:

(i) G has at most a finite number of conjugacy classes of rank one and rank two parabolic subgroups (Sullivan; see [Feighn and McCullough 1987]).

(ii) G has at most a finite number of conjugacy classes of finite subgroups [Feighn and Mess 1991].

Hint: For (i), use the compact core or the relative compact core and the fact that, corresponding to the rank one and rank two cusps, there are mutually disjoint cusp cylinders and cusp tori with the property that a simple nontrivial loop on one is not freely homotopic to one on another [Kulkarni and Shalen 1989]. For the second, first apply Selberg's Lemma (page 68) to get a torsion-free, normal subgroup of finite index H. The finite group $F = G/H$ is isomorphic to a group of automorphisms of $\mathcal{M}(H)$. It is shown in [Feighn and Mess 1991] that one can choose a compact core C of $\mathcal{M}(H)$ to be invariant under F; for this, C/F is compact.

3-16. *Retractions.* [Epstein and Marden 1987]Let K be a hyperbolically convex set in $\mathbb{H}^3$. The retraction map $r : \mathbb{H}^3 \setminus K \to K$ is defined as follows. For each $\vec{x} \in \mathbb{H}^3 \setminus K$, $r(\vec{x})$ is to be that point of K closest, in the hyperbolic metric, to $\vec{x}$. This closed point is uniquely attained. In the hyperbolic metric $d(\cdot\,,\cdot)$, show that the map r is Lipschitz: $d(r(\vec{x}), r(\vec{y})) < d(\vec{x}, \vec{y})$.

Hint: Normalize so that the geodesic from $r(\vec{x})$ to $r(\vec{y})$ lies on the vertical axis ℓ in the upper half-space model. Draw the planes orthogonal to ℓ through $r(\vec{x})$ and $r(\vec{y})$. The geodesic segment between $r(\vec{x})$ and $r(\vec{y})$ lies in K. The points $\vec{x}$, $\vec{y}$ cannot lie in the open set bounded by the two planes.

3-17. *Cylindrical manifolds.* Suppose G is geometrically finite and $\partial\mathcal{M}(G)$ is incompressible. Let $C \subset \mathcal{M}(G)$ be an essential cylinder; $\mathcal{M}(G) \setminus C$ has one or two components M_1, M_2. Choose one of these, say M_1, and consider a lift $M_1^* \subset \mathbb{H}^3$. Set $G_1 = \mathrm{Stab}(M_1^*)$. Describe $\Omega(G_1)$ in terms of $\Omega(G)$ and the topological type of $\partial\mathcal{M}(G_1)$ in terms of $\partial\mathcal{M}(G)$. In turn cut $\mathcal{M}(G_1)$ along an essential cylinder, if it has one. Show that this process must end after a finite number of steps. Classify the different possibilities you can end up with.

A parabolic $T \in G$ is called *accidental* if there is a component $\Omega \subset \Omega(G)$ such that $T(\Omega) = \Omega$ in which T has the three (equivalent) properties: (i) T has no horodisk in Ω; (ii) T corresponds to a loxodromic transformation in the Riemann map image or the universal cover $\mathbb{H}^2$ of Ω; (iii) the fixed point of T lies in the impression of two distinct prime ends of $\partial\Omega$. The simplest example is the transformation $z \mapsto z + 1$ acting in the strip $S = \{z : 0 < \mathrm{Im}\, z < \pi\}$. The Riemann map $z \mapsto e^z$ maps S unto the upper half-plane. The parabolic T is transferred to the loxodromic $w \mapsto ew$. In truth the attribute "accidental" is singularly inappropriate, as there is nothing accidental about the appearance of an accidental parabolic.

From the three-dimensional point of view, it is possible that an "essential cylinder" just bounds a solid cusp tube for a rank one cusp. Then the result of cutting as proposed above does not change the group and indeed we have decided that such a cylinder is not officially called an essential cylinder. However it is entirely possible that one of the boundary components of an essential cylinder be retractable to a puncture, and the other not. This is exactly the situation of an "accidental" parabolic in a geometrically finite manifold. Suppose you only cut the manifold along essential cylinders associated with such parabolics. Show that after a finite number of steps you will end up with a group or groups that no longer have such any such parabolics: every cyclic parabolic group pairs exactly two punctures and is not represented by any homotopically different simple loop in the boundary; see [Abikoff and Maskit 1977].

There is no reason to believe that a core curve c of an essential cylinder C is primitive; if $A \in G$ is determined by c, is it possible that $A = B^n$ for $B \in G$ and $n > 1$?

3-18. *Conical limit points.* If G is a nonelementary kleinian group, a point $\zeta \in \Lambda(G)$ is called a *conical limit point* if the following is true. Let $\gamma(t)$, $0 \le t < \infty$, be a geodesic ray ending at ζ. Given $O \in \mathbb{H}^3$, there exists $r > 0$ such that there is an infinite subsequence of the orbit $G(O)$ that lies in the r-tubular neighborhood about γ (and hence converges to ζ).

In the quotient manifold, the condition means that the projection of the ray $\gamma(t)$ is *recurrent* in the sense that given any point $\pi(O) \in \mathbb{H}^3/G$ there is an infinite sequence $t_n \to \infty$ such that each $\pi(\gamma(t_n))$ is within distance r of $\pi(O)$. Put another way, there exists a compact subset $K \subset \mathbb{H}^3/G$ such that $\pi(\gamma(t))$ intersects K infinitely many times. A loxodromic fixed point is always a conical limit point, and a parabolic fixed point is never one.

If $\mathcal{M}(G)$ is geometrically finite, and ζ is not a parabolic fixed point, then $\pi(\gamma(t))$ will lie in a compact set because it cannot asymptotically penetrate the universal horoballs.

Prove that G is geometrically finite if and only if all limit points except parabolic fixed points are conical limit points [Beardon and Maskit 1974]. *Hint:* All geodesics lie in the convex hull of $\Lambda(G)$.

3-19. *Quasiisometries.* A *quasiisometry* of $\mathbb{H}^3$ (or of any $\mathbb{H}^n$) is a map $f : \mathbb{H}^3 \to \mathbb{H}^3$ that satisfies

$$\frac{1}{L} d(x, y) - a \le d(f(x), f(y)) \le L d(x, y) + a \tag{3.14}$$

for some $L \ge 1$ and $a \ge 0$. The map f need not be a homeomorphism nor even continuous, just asymptotically Lipschitz. It is called a *Lipschitz map* if the right inequality holds for $a = 0$. The minimum factor L is called the *Lipschitz constant* for f. Initially Mostow used “pseudo-isometries”, which satisfy (3.14) except on the right side $a = 0$ so the map is Lipschitz; the long range properties are the same whether or not $a = 0$. A homeomorphism f is *L-bilipschitz* if (3.14) holds with $a = 0$. An L-bilipschitz map on $\mathbb{H}^2$ or $\mathbb{H}^3$ is L^2-quasiconformal. The converse is not true in general (for an example, consider the radial stretch $\vec{x} \mapsto |\vec{x}|^\alpha \vec{x}$, $-1 < \alpha < 0$).

An equivalent definition is perhaps more illuminating: there exists constants $K \geq 1$ and $d_0 \geq 0$ such that

$$\begin{aligned} K^{-1} d(x, y) &\leq d(f(x), f(y)) \quad \text{for all } x, y \in \mathbb{H}^3 \text{ with } d(x, y) \geq d_0, \\ K\, d(x, y) &\geq d(f(x), f(y)) \quad \text{for all } x, y \in \mathbb{H}^3. \end{aligned}$$

For example, $\mathbb{H}^2$ and $\mathbb{H}^3$ are quasiisometric to the Cayley graph dual to the tessellation by a fundamental polygon or polyhedron for a fuchsian group or a kleinian group whose respective quotients are closed [Cannon and Cooper 1992]: The graphs look like $\mathbb{H}^2$ and $\mathbb{H}^3$ if you look at them from afar. In fact, a hyperbolic group (Exercise 2-17) is quasiisometric to $\mathbb{H}^2$ or $\mathbb{H}^3$ if and only if it is a fuchsian group representing a closed surface [Boileau et al. 2003, Theorem 6.18] or a kleinian group representing a closed manifold [Cannon and Cooper 1992].

Using Equation (8.27) of Exercise 8-9, prove that quasiisometries have the following properties [Efremovič and Tihomirova 1964; Thurston 1979, p. 5.39]:

(i) If γ is a geodesic ray to a point $\zeta \in \partial\mathbb{H}^3$, then $f(\gamma)$ has a well defined end point on $\partial\mathbb{H}^3$. Denote the end point by $f(\zeta)$.
(ii) There exists a constant $M < \infty$, such that for any $x \in \gamma$, $d(f(x), \gamma') < M$, where γ' denotes a geodesic ray ending at $f(\zeta)$.
(iii) The extension of f to $\partial\mathbb{H}^3$ is a homeomorphism.
(iv) If $f(\zeta) = \zeta$ for all $\zeta \in \partial\mathbb{H}^3$, then $\sup_{x\in\mathbb{H}^3} d(x, f(x)) < \infty$.
(v) More generally, if f_1, f_2 are quasiisometries with the same boundary values, there exists a constant $B < \infty$ such that $d(f_1(x), f_2(x)) < B$ for all $x \in \mathbb{H}^3$.

An additional important property is that the extension to $\partial\mathbb{H}^3$ is quasiconformal [De-Spiller 1970]. To prove this it is necessary to show the *metric definition* of quasiconformality is verified: Let τ denote the spherical metric on $\partial\mathbb{H}^3$. Set

$$\begin{aligned} L(f, r)(\zeta) &= \sup_{\tau(\zeta',\zeta)=r} \tau(f(\zeta'), f(\zeta)), \\ l(f, r)(\zeta) &= \inf_{\tau(\zeta',\zeta)=r} \tau(f(\zeta'), f(\zeta)), \\ H(f, \zeta) &= \overline{\lim}_{r\to 0} \frac{L(f, r)(\zeta)}{l(f, r)(\zeta)}. \end{aligned}$$

The restriction of f to $\partial\mathbb{H}^3$ is quasiconformal if there exists $K^* < \infty$ such that $H(f, \zeta) < K^*$ for all $\zeta \in \partial\mathbb{H}^3$.

The use of this theory to prove Mostow's Rigidity Theorem is indicated at the end of Section 3.12. It is perhaps interesting to digress to summarize the history of Mostow's result. His original announcement and proof had the hypothesis that there is a quasiconformal mapping between closed manifolds $f : \mathcal{M}(G) \to \mathcal{M}(H)$ (actually in n-dimensions). Most of his 1968 paper was devoted to the proof, following earlier work of Gehring, that f, when lifted to $\mathbb{H}^n$, can be extended to $\partial\mathbb{H}^n$ and is there a quasiconformal mapping. He then used properties of quasiconformal mappings together with some ergodic theory to prove rigidity. Ahlfors immediately recognized the importance of Mostow's result and worked to make the proof more transparent to complex analysts. In a short unpublished manuscript, which assumed the boundary extension property which was known to him, Ahlfors simplified Mostow's proof, using fewer properties of quasiconformal mappings and less ergodic theory. The book [Mostow 1973] contains an entirely new proof of a more general theorem. There Mostow introduced the notion of a pseudo-isometry (now called a quasiisometry) and developed its properties. While Mostow was working on his generalization, G. A. Margulis independently published in 1970 a page-and-a-half "plan of a proof" of a less encompassing generalization of Mostow's original theorem. He too used a method akin to quasiisometries. The key feature that extension to the boundary is quasiconformal which was published in its own right in [De-Spiller 1970] was thus apparently independently discovered by Mostow and Margulis in the course of their application.

3-20. *Hausdorff dimension.* The notion of Hausdorff dimension is used to measure the "size" of point sets with smooth curves having dimension one and isolated points having dimension zero. The *α-dimensional Hausdorff measure* of a closed set $X \subset \mathbb{C}$ (or more generally, of a Borel set) is defined in terms of

$$\Lambda_\alpha(X) = \lim_{\varepsilon\to 0}\left(\inf_{\{D_k\}}\sum_k \operatorname{Diam}(D_k)^\alpha\right),$$

where the infimum is taken over all covers $\{D_k\}$ of X by euclidean disks of diameters at most ε. The *Hausdorff dimension* is defined as

$$\dim\ X = \inf\{\alpha : \Lambda_\alpha(X) = 0\}.$$

The inequality

$$\sum_k (\operatorname{Diam} D_k)^\beta \le \varepsilon^{\beta-\alpha}\sum_k (\operatorname{Diam} D_k)^\alpha,$$

which implies that $\Lambda_\beta(X) \le \varepsilon^{\beta-\alpha}\Lambda_\alpha(X)$, shows that $\Lambda_\alpha(X) = 0$ if $\alpha > \dim\ X$ while $\Lambda_\alpha(X) = \infty$ if $\alpha < \dim\ X$.

If f is an L-bilipschitz map of X, we have

$$L^{-\alpha}\Lambda_\alpha(X) \le \Lambda_\alpha(f(X)) \le L^\alpha\Lambda_\alpha(X).$$

For sets $X \subset \mathbb{C}$, $0 \le \dim\ X \le 2$. A connected closed set without interior which has Hausdorff dimension > 1 is called a *fractal*. Upper estimates of the Hausdorff dimension are often found by using a special covering and by the following estimate. Assume X is a bounded set. Let $N(\varepsilon)$ denote the minimum number of round disks of diameter ε needed to cover X. Then

$$\dim\ X \le \liminf_{\varepsilon\to 0}\frac{\log N(\varepsilon)}{-\log\varepsilon}.$$

In this estimate, a square grid of side length ε covering X and such that each square intersects X can replace the minimal cover by disks. For a careful development of the theory for plane point sets from the point of view of conformal mapping see [Pommerenke 1992].

Thanks to the fundamental paper [Bishop and Jones 1997], added to earlier results (see the discussion in [Canary and Taylor 1994]), we can assert:

Theorem 3.14.1. *Suppose G is a finitely generated kleinian group.*

(*i*) $\dim(\Lambda(G)) > 0$ *if and only if G is nonelementary.*
(*ii*) $\dim(\Lambda(G)) < 2$ *if and only if G is geometrically finite.*
(*iii*) $\dim(\Lambda(G)) > 1$ *if $\Lambda(G) \subset \mathbb{S}^2$ is connected but is not a circle (in which case* $\dim(\Lambda(G)) = 1$).
(*iv*) *If* $\dim(\Lambda(G)) < 1$ *then $\Lambda(G)$ is totally disconnected.*

In particular for a Schottky group G, $\Lambda(G)$ has zero area but positive Hausdorff dimension, in Mandelbrot's terminology, it is "fractal dust". At the other extreme, for the singly degenerate groups of Section 5.8, in which $\Omega(G)$ is connected and simply connected, we have $\dim \Lambda(G) = 2$. For further information see the excellent survey [Matsuzaki and Taniguchi 1998].

Another measure for a kleinian group G acting on the ball model of $\mathbb{H}^3$ is its *critical exponent* which is defined as

$$\delta(G) = \inf\left\{s : \sum_{g\in G} e^{-sd(0,g(0))} < \infty\right\} = \inf\left\{s : \sum_{g\in G}\left(\frac{1-|g(0)|}{1+|g(0)|}\right)^s < \infty\right\},$$

where 0 is the center of the ball and $d(\cdot\,,\cdot)$ is hyperbolic distance. Use Exercise 1-14 to show that

$$\delta(G) = \inf\left\{s : \sum_{g\in G}(1-|g(0)|)^s < \infty\right\} = \inf\{s : \sum_{g\in G}|g'(0)|^s < \infty\}.$$

We will see in Exercise 3-22 that $\delta(G) \le 2$. Moreover

$$\delta(G) = \dim \Lambda_c(G) \le \dim\ \Lambda(G),$$

where $\Lambda_c(G)$ denotes the set of conical limit points (see Exercise 3-18).

It is amazing that to know the critical exponent is to know the Hausdorff dimension. Combining the solution of Ahlfors' conjecture with a result in [Bishop and Jones 1997] yields:

Theorem 3.14.2. *If G is a finitely generated group with $\Omega(G) \ne \varnothing$, then*

$$\delta(G) = \dim(\Lambda(G)).$$

If $\Omega(G) = \varnothing$, both sides have the value 2.

3-21. This problem and the next are the first steps into the study of the ergodic theory of kleinian groups. See [Nicholls 1989] for an introduction to the theory, especially the construction of invariant measures on the limit set. Also see [Patterson 1987].

Let G be a discrete group acting in the ball model. Define the orbital counting function to be

$$N(r;\vec{x},\vec{y}) = \mathrm{card}\{g \in G : d(\vec{x}, g(\vec{y})) < r\}.$$

Prove that there is a constant C depending on G and $\vec{y}$, such that for any point $\vec{x}$,

$$N(r,\vec{x},\vec{y}) \ < \ Ce^{2r}.$$

If in addition G has a fundamental polyhedron of finite volume, show that there is a constant $C_0 = C_0(G,\vec{x},\vec{y})$ such that

$$N(r;\vec{x},\vec{y}) \ > \ C_0 e^{2r} \quad \text{for all large } r.$$

Hint: Prove first that

$$\mathrm{Vol}(\{\vec{x} : d(0,\vec{x}) < r\}) = 2\pi \int_0^r \sinh^2(t)dt,$$

where $t = |\vec{x}|$ and $d(0,\vec{x}) = \log \frac{1+t}{1-t}$. Find $\varepsilon > 0$ such that no two elements of the G-orbit of the ball $B_\varepsilon(\vec{y})$ of radius ε centered at $\vec{y}$ overlap. However if $\vec{y}$ is an elliptic fixed point of order m, then for each element of the orbit, m-images will coincide. Assuming $\vec{y}$ is not a fixed point, show that

$$\mathrm{Vol}(B_\varepsilon(\vec{y}))N(r;\vec{x},\vec{y}) < \mathrm{Vol}(B_{r+\varepsilon}(\vec{x})) = 2\pi \int_0^{r+\varepsilon} \sinh^2(t)dt < \frac{1}{4}\int_0^{r+\varepsilon} e^{2t}dt.$$

Determine the corresponding statements for the action of a fuchsian group in the unit disk model of $\mathbb{H}^2$.

3-22. Suppose G is a discrete group. Refer back to Exercise 1-14 and prove that

$$\sum_{g\in G} e^{-\alpha d(0,g(0))} < \infty \quad \text{for all } \alpha > 2.$$

Hint: Consider Exercise 3-21 and

$$\sum_{g\in G: d(0,g(0))<r} e^{-\alpha d(0,g(0))} = \int_0^r e^{-\alpha t} dN(t;0,0)$$
$$= N(r;0,0)e^{-\alpha r} + \alpha \int_0^r N(t;0,0)e^{-\alpha t}dt.$$

The same expression for the disk model of $\mathbb{H}^2$ converges for $\alpha > 1$.

Let $\delta(G)$ be the critical exponent for G as defined in Exercise 3-20. We know from Exercise 3-20 that $\delta(G) \le 2$ for groups acting in the ball model and that $\delta(G) \le 1$ for fuchsian groups acting in the unit disk. The group is said to be of *convergence type* if

$$\sum_{g\in G} e^{-\delta(G)d(0,g(0))} \ < \ \infty;$$

otherwise G is said to be of *divergence type*.

Prove that if G has a fundamental polyhedron of finite volume in the ball model, or a fundamental polygon of finite area in the disk model, then G is of divergence type. If however the limit set $\Lambda(G) \neq \mathbb{S}^2$ in the ball model or $\neq \mathbb{S}^1$ in the disk model, then G is of convergence type.

Hint: Let D be a closed disk in the ordinary set $\Omega(G)$. We can assume that the elements of its G-orbit are mutually disjoint. Let $d\sigma$ denote the area form on $\mathbb{S}^2$: in spherical coordinates, $d\sigma = \sin\varphi\, d\varphi\, d\theta$. Then

$$4\pi > \sum_{g\in G} \operatorname{Area}(g(D)) = \sum_{g\in G} \iint_D |g'(w)|^2\, d\sigma.$$

To finish, deduce from Exercise 1-12 that when $|w| = |g(w)| = 1$,

$$|g'(w)||g(w) - g(0)|^2 = 1 - |g(0)|^2, \quad 1 - |g(0)| < 4|g'(w)|.$$

On the other hand, if G has a polyhedron with finite volume, write again

$$\begin{aligned}\sum_{g\in G: d(0,g(0))<r} e^{-2d(0,g(0))} &= \int_0^r e^{-2t}\, dN(t;0,0)\\ &= N(r;0,0)e^{-2r} + 2\int_0^r N(t;0,0)e^{-2t}\, dt.\end{aligned}$$

Apply Exercise 3-21 to finish the job.

The following result appears as [Matsuzaki and Taniguchi 1998, Theorem 5.15] and incorporates results from [Ahlfors 1981] and [Sullivan 1981]. For the proof, see [Nicholls 1989].

Theorem 3.14.3. *The following statements about a kleinian group G are equivalent*:

(i) *The conical limit set $\Lambda_c(G)$ has Lebesgue measure 4π on $\mathbb{S}^2$.*
(ii) *G is of divergence type.*
(iii) *$\mathcal{M}(G)$ does not support a hyperbolic Green's function.*
(iv) *G acts ergodically on $\mathbb{S}^2 \times \mathbb{S}^2$.*
(v) *The geodesic flow on the unit tangent bundle of Int$(\mathcal{M}(G))$ is ergodic.*

The group G is said to act *ergodically* on $\mathbb{S}^2$, or on $\mathbb{S}^2 \times \mathbb{S}^2$, if and only if the following holds: Given a measurable set X invariant under the action of G, the Lebesgue measure of either X or of the complement of X vanishes. Here the action of $g \in G$ on $\mathbb{S}^2 \times \mathbb{S}^2$ is $(x, y) \mapsto (g(x), g(y))$. Thus ergodic action on the product implies ergodic action on $\mathbb{S}^2$ itself. For a discussion of hyperbolically harmonic functions and their boundary values see Exercise 5-1.

3-23. *Poincaré series.* Suppose G is a fuchsian group of convergence type acting in the unit disk $\mathbb{D}$. Suppose $f(z)$ is a bounded analytic function in $\mathbb{D}$. Prove that the *Poincaré series*

$$\Phi(z) = \sum_{g\in G} f(g(z))g'(z)$$

is an analytic function in $\mathbb{D}$ that satisfies the functional relation

$$\Phi(g(z))g'(z) = \Phi(z) \quad \text{for all } g \in G,\ z \in \mathbb{D}.$$

That is, $\Phi(z)\,dz$ is an invariant form under G. It projects to a *holomorphic differential* on the quotient Riemann surface $\mathbb{D}/G$.

Now suppose instead that G is of divergence type. Prove that the Poincaré series

$$\Phi(z) = \sum_{g \in G} f(g(z))(A')^2(z)$$

is analytic in $\mathbb{D}$ and satisfies

$$\Phi(g(z))(g')^2(z) = \Phi(z) \quad \text{for all } g \in G,\ z \in \mathbb{D}.$$

The invariant form $\Phi(z)\,dz^2$ projects to an *holomorphic quadratic differential* on $\mathbb{D}/G$.

3-24. *Isotopy.* Two mappings between manifolds $f, g : M \to N$ are said to be *homotopic* if there is a continuous flow $F_t : M \to N,\ 0 \le t \le 1$ such that $F_0 = f$, $F_1 = g$. In contrast f, g are *isotopic* if f, g are homeomorphisms and at time $t,\ 0 \le t \le 1$, F_t is a homeomorphism. Show that a homeomorphism of a 3-manifold onto itself can be homotopic but not isotopic to the identity. (*Hint:* $S \times [0, 1]$ flipped over. A homeomorphism which on the boundary is a Dehn twist (Example 5-11) about the boundary of a compressing disk.) In contrast, on a surface, homeomorphisms which are homotopic are also isotopic.

Two simple curves γ_1, γ_2 in a surface S are said to be *isotopic* if there is a continuous map $F_t : \mathbb{S}^1 \times [0, 1] \to S$ such that for each t. F_t is a homeomorphism of $\mathbb{S}^1$ into S such that F_0 gives γ_1 and F_1 gives γ_2. Two simple curves that are freely homotopic in S are also isotopic.

On the other hand, the study of knots in $\mathbb{S}^3$ rests on the difference between isotopy and homotopy: Any knot is homotopic to an embedded circle, but is isotopic to one if and only if its complement is homeomorphic to the complement of an embedded circle.

3-25. *Homotopic isometries.* Prove that homotopic *isometries* of a hyperbolic manifold are identical, provided the fundamental group is nonabelian. *Hint:* lift to $\mathbb{H}^3$.

3-26. *Voronoi diagrams, Delaunay triangulations, and polyhedra.* Given a discrete set of points X in $\mathbb{H}^3$ (or $\mathbb{H}^n$), make the following construction. Given $x \in X$ construct the cell C_x with center x consisting of all points closer to x than to any other point in X. It is the intersection of all half-spaces containing x which are bounded by the planes orthogonal to the line segments between x and the other points of X. The Voronoi diagram consists of the totality of cells built around elements of X. It is a subdivision of $\mathbb{H}^3$. Each face is shared by two cells, and each vertex is shared by at least three cells. If v is a vertex, then there is a sphere σ_v about v which contains the centers of all those cells sharing the vertex v. Moreover the ball bounded by σ_v lies in the union of the cells sharing v and its interior contains no points of X.

The Delaunay triangulation is dual to the Voronoi diagram. Given $x \in X$ draw a geodesic segment from x to the points of X whose cells share a face with C_x, and continue this process for all elements of X. We obtain a decomposition of space into polyhedra. There is one polyhedron P_v for each vertex v; the edges of P_v are the line segments between the centers of the cells that share the vertex v. The vertices of P_v are the centers x of these cells. The totality of polyhedra $\{P_v\}$ are the Delaunay "triangles" (the term comes from the 2-dimensional case).

Efficient ways of numerically finding Voronoi diagrams and Delaunay triangulations is an important issue in computer science.

If X consists of the orbit of x under a discrete group G without a fixed point at x, C_x is precisely the Dirichlet region with basepoint x. The dual Delaunay "triangles" give a dual G-invariant decomposition of $\mathbb{H}^3$ by polyhedra.

There is an interesting limiting case. Suppose $\mathcal{M}(G)$ is a geometrically finite manifold of finite volume. In the upper half-space model, say, assume ∞ is a parabolic fixed point. Construct the Ford "polyhedron" $\mathcal{F}$ with "center" ∞. As we have seen on page 120, $\mathcal{F}$ is invariant under the stabilizer Stab_∞ of ∞. Its orbit under the cosets of Stab_∞, is a Voronoi diagram. To obtain the Delaunay triangulation, draw the geodesics between ∞ and the centers of the polyhedra with share faces with $\mathcal{F}$, and so on. Show that there results a tessellation of $\mathbb{H}^3$ by ideal polyhedra centered on the interior vertices ζ of $\mathcal{F}$ and its orbit. Down below, there is a decomposition of $\mathcal{M}(G)$ into a finite number of ideal polyhedra. For more discussion see [Weeks 1993; Petronio and Weeks 2000].

3-27. What is the maximum and minimum number of sides that a Dirichlet region can have for a closed hyperbolic surface of genus $g \geq 0$ with $b \geq 0$ punctures?

The square once-punctured torus is defined by the property that there are two geodesics of equal length that cross once. The hexagonal once-punctured torus has the property that there are three distinct geodesics of equal length that intersect at a point. Is it true that the hexagonal torus has the longest shortest geodesic in its deformation space? Construct the corresponding symmetric Dirichlet regions and determine the generating matrices.

3-28. (V. Markovic) In the ball model, suppose $\Omega \subset \mathbb{S}^2$ is a simply connected component of $\Omega(G)$, for some nonelementary group G. Take the Dirichlet region $\mathcal{P}_p$ centered at a point $p \in \mathrm{Dome}(\Omega)$. Then let p approach a point on $\partial\mathrm{Dome}(\Omega) \subset \Lambda(G)$. Show that the euclidean diameter of $\mathcal{P}_p$ tends to zero. In fact if the euclidean distance of p from $\mathbb{S}^2$ is ε, then the euclidean diameter of $\mathcal{P}_p$ is $\leq \sqrt{\varepsilon}$.

3-29. *Isomorphisms that determine homeomorphisms* [Tukia 1985b]. Suppose that $\varphi : G \to H$ is an isomorphism between geometrically finite, nonelementary groups without elliptics such that $\varphi(G)$ is parabolic if and only if $g \in G$ is so. The problem is to determine when φ is induced by a homeomorphism. Start by showing that there is a uniquely determined homeomorphism f_φ between the respective limit sets which induces φ. The map f_φ sends the attracting loxodromic fixed point of $g \in G$ to the attracting fixed point of $\varphi(g)$. There is also need for [Tukia 1985b, Lemma 3.4], which says there is a quasiisometry F of the convex hull $F : \mathcal{C}(G)$ into $\mathcal{C}(H)$ which induces φ and with the property that in the hyperbolic distance, $d(x, F(\mathcal{C}(G)))$ is uniformly bounded for $x \in \mathcal{C}(H)$. This is akin to one of the techniques used for Mostow's theorem.

There is a celebrated theorem of Fenchel and Nielsen concerning isomorphisms $\varphi : \Gamma \to \Gamma'$ between two fuchsian groups. Namely φ is induced by an orientation preserving or reversing homeomorphism $\mathbb{H}^2 \to \mathbb{H}^2$ if and only if the following property holds: The axes of loxodromics $g, h \in \Gamma$ intersect in $\mathbb{H}^2$ if and only if the images $\varphi(g), \varphi(h)$ are also loxodromic and have intersecting axes.

To generalize this we say that a loxodromic $g \in G$ and a quasifuchsian subgroup $\Gamma \subset G$ intersect provided the fixed points of g lie in $\Omega(\Gamma)$, one in each component. Suppose $\varphi : G \to H$ is an isomorphism between two geometrically finite groups without parabolics or elliptics, and φ preserves intersection in the sense that a loxodromic $g \in G$ and quasifuchsian $\Gamma \subset G$ intersect if and only if $\varphi(g)$ and $\varphi(\Gamma)$ intersect. Here $\varphi(\Gamma)$ is necessarily quasifuchsian because $f_\varphi(\Lambda(\Gamma))$ is a topological circle.

Suppose G is geometrically finite but not quasifuchsian. Prove that if Ω is a component of $\Omega(G)$ there is a uniquely determined component Ω' of $\Omega(H)$ such that $f_\varphi(\partial\Omega) = \partial(\Omega')$ and $\varphi(\mathrm{Stab}_\Omega) = \mathrm{Stab}_{\Omega'}$. The intersection property comes in to establish that if $\sigma \subset \Lambda(G)$ is a the limit set of a quasifuchsian subgroup of G, then $x, y \in \Lambda(G) \setminus \sigma$ lie in different components of $\mathbb{S}^2 \setminus \sigma$ if and only if $f_\varphi(x)$, $f_\varphi(y)$ are in different components of $\mathbb{S}^2 \setminus f_\varphi(\sigma)$. The bottom line is:

Theorem 3.14.4 [Tukia 1985b, Theorem 4.7]. *Suppose $\varphi : G \to H$ is an isomorphism between nonelementary, geometrically finite groups without elliptics such that $\varphi(g)$ is parabolic if and only if $g \in G$ is so. If $\Lambda(G)$ is connected, assume that φ preserves intersections. If $\Lambda(G)$ is not connected, certain orientability conditions must be satisfied for quasifuchsian subgroups and rank two parabolics. Then φ is induced by a quasiconformal homeomorphism Φ of $\mathbb{H}^3 \cup \mathbb{S}^2$.*

The orientability condition for a quasifuchsian subgroup $H \subset G$ is that the map f_φ restricted to $\Lambda(H)$ can be extended to an orientation preserving map of $\mathbb{S}^2$ which sends attracting fixed points of loxodromics in H to attracting fixed points of their φ-images. For a rank two parabolic subgroup, φ needs to be induced by an orientation preserving map of $\mathbb{S}^2$.

3-30. *Intersections.* If G_1, G_2 are finitely generated fuchsian groups, prove that the intersection $G_1 \cap G_2$ is also finitely generated.

If H is a finitely generated subgroup of the fuchsian group G and the limit sets are the same, then H has finite index in G.

Both these results can be found in [Greenberg 1960].

If G_1 and G_2 are finitely generated subgroups of the not necessarily finitely generated H with $\Omega(H) \neq \varnothing$, then

$$\Lambda(G_1) \cap \Lambda(G_2) = \Lambda(G_1 \cap G_2).$$

This is proved in [Anderson 1996].

3-31. [Greenberg 1977, Theorem 2.5.8] Suppose $\alpha : z \mapsto z + 1$ is a generator of a rank one parabolic subgroup of the finitely generated kleinian group G. Note that for any $g \in G$, $\mathcal{O} \in \mathbb{H}^3$, the perpendicular bisector of $[g(\mathcal{O}), \alpha g(\mathcal{O})]$ is a vertical plane. Show that the fundamental polyhedron $\mathcal{P}_{\mathcal{O}}$ lies in a slab $\{(z = x + iy, t) : a \leq x \leq a + 1\}$. There is a universal horoball at ∞. Suppose further that α has a horodisk $\mathcal{H} = \{z : y > b\} \subset \Omega(G)$. If for the euclidean closure $H = \overline{\mathcal{P}}_{\mathcal{O}} \cap \mathcal{H} \neq \emptyset$, show that $H = \{z : a \leq c_1 \leq x \leq c_2 \leq a + 1\}$ for some c_1, c_2. From this prove MacMillan's theorem that $\overline{\mathcal{P}}_{\mathcal{O}} \cap \Omega(G)$ has a finite number of sides.

3-32. *Earthquakes.* This is to introduce Thurston's theory of earthquakes [1986a]. For this purpose let ℓ be the positive imaginary axis in the upper half-plane model UHP of $\mathbb{H}^2$. Denote the left and right quarter planes determined by ℓ by A and B; A and B have orientations inherited from $\mathbb{C}$. From the point of view of A, a *left earthquake* with fracture line ℓ is a discontinuous map which fixes A pointwise, and in B is an isometry moving B *to the left* with respect to A; that is, it moves B in the positive direction with respect to the positive orientation of ∂A. Therefore in B it has the form $z \mapsto kz$, for $k > 1$. It is uniquely determined once the displacement along ℓ is dictated.

If instead we require that B be fixed, the left earthquake along ℓ moves A to the *left* from the point of view of a person standing in B. In A it has the form $z \mapsto k^{-1}z$.

Next suppose we have a finite lamination. Fix a gap σ as the base of operations. Suppose μ is a positive transverse measure—that is, to each leaf of the lamination is assigned a positive number as atomic measure. The earthquake will be the identity on sigma. A transverse geodesic based in σ will cross a number of leaves. Carry out a sequence of left earthquakes in sequence along the various leaves, using the displacement assigned by μ.

Here is a more formal definition. Suppose $\Lambda \subset \mathbb{H}^2$ is a geodesic lamination. A *left earthquake* is a possibly discontinuous injective and surjective map $E : \mathbb{H}^2 \to \mathbb{H}^2$ which is an isometry on each leaf of Λ and on each complementary component. Given two gaps and/or leaves $X \neq Y$, a line ℓ is said to be weakly separating if any path from a point of X to a point of Y intersects ℓ. Let E_X, E_Y denote the respective isometric restrictions of E. We require that the *comparison isometry* $E_X^{-1} \circ E_Y$ be loxodromic, that its axis ℓ weakly separate X and Y, and that it translate to the left, when viewed from X. This last requirement means that the direction of translation along ℓ agrees with the orientation induced from $X \subset \mathbb{H}^2 \setminus \ell$. The case that one of X, Y is a line in the boundary of the other is exceptional in that the comparison map is the identity.

The earthquake maps Λ to another lamination Λ'. The inverse of a left earthquake is a right one.

If Λ has a finite number of leaves, left earthquakes are constructed as illustrated above. Thurston proves that these finite earthquakes are dense in all left earthquakes, in the topology of uniform convergence on compact sets.

A left earthquake between two Riemann surfaces is an injective, surjective map which lifts to a left earthquake of $\mathbb{H}^2$. In particular Λ is invariant under the deck transformations. However if one or more leaves of Λ project to simple geodesics, lifts are determined only up to "twists" along the geodesics. To avoid this ambiguity one can associate the earthquake with the homotopy type of a homeomorphism between the surfaces. A more common way, is to start with both an invariant lamination in $\mathbb{H}^2$, and an invariant transverse measure (more of this below).

Earthquake Theorem [Thurston 1986a]. *Every continuous orientation preserving map $\partial\mathbb{H}^2 \to \partial\mathbb{H}^2$ is the boundary values of a left earthquake E of $\mathbb{H}^2$. The lamination Λ is uniquely determined. On Λ, E is uniquely determined except along those leaves ℓ on which it is discontinuous. For each such ℓ, there is a range of choices of translations ranging between the limiting values of E on the two sides; all the choices have the same image in $\mathbb{H}^2$.*

Suppose $R_i = \mathbb{H}^2/G_i$, $i = 1, 2$, are arbitrary Riemann surfaces with possible boundary contours ∂R_i coming from the action of G_i on maximal open intervals of discontinuity on $\partial\mathbb{H}^2$. Assume $h : R_1 \to R_2$ is an (orientation preserving) homeomorphism which extends to a continuous map $\partial R_1 \to \partial R_2$. Then the boundary values on $\partial\mathbb{H}^2$ of a lift of h are the boundary values of a left earthquake of $\mathbb{H}^2$ which projects back to a left earthquake $E : \mathbb{R}_1 \to R_2$. Moreover, E has the same uniqueness indicated above.

This is a very general theorem. The second statement (which includes the first) follows from the first as lifts of h extend to continuous maps of $\partial\mathbb{H}^2$. Punctures on R_2 do not necessarily come from punctures on R_1.

Associated to any left earthquake is a nonnegative transverse Borel measure μ. Two earthquakes corresponding to the same (Λ, μ) have isometric images. The measure is constructed by a process akin to Riemann integration (see [Epstein and Marden 1987]).

Normally one only works with the restricted class of *uniformly (locally) bounded* earthquakes. These are the class of earthquakes whose transverse measures have the property that for some $K < \infty$, $\mu(\tau) < K$ for all transverse geodesic segments τ of unit length.

The boundary values on $\partial\mathbb{H}^2$ of uniformly locally bounded earthquakes are quasisymmetric (that is, 1-quasiconformal) homeomorphisms, which means their boundary values have quasiconformal extensions to $\mathbb{H}^2$ (and which are equivariant if (Λ, μ) is invariant under deck transformations). In the other direction, the boundary values of a quasiconformal mapping $\mathbb{H}^2 \to \mathbb{H}^2$ (say the lift of a map between surfaces) are also the boundary values of a uniformly bounded left earthquake as described in the Earthquake Theorem. In particular this is true for bounded earthquakes and quasiconformal mappings between Riemann surfaces of finite area. For the details, consult [Thurston 1986a].

Given (Λ, μ), an *earthquake flow* is the earthquake E_t associated with $(\Lambda, t\mu)$ with $0 \le t$. For an application of this technique to the solution of the Nielsen Realization Problem, see [Kerckhoff 1983].

In summary, in the dictionary entry relating geometry to complex analysis, earthquakes are the analogue of quasiconformal mappings used to deform conformal structure.

3-33. *The Nielsen kernel.* Suppose G is a fuchsian group in the unit disk $\mathbb{D}$ with $R = \mathbb{D}/G$ the interior of a compact, bordered Riemann surface $\overline{R}$ of genus $g \geq 0$, $n \geq 0$ punctures and $m \geq 1$ boundary contours. There is a set $C_1, \ldots, C_m$ of mutually disjoint simple geodesics such that C_i bounds an annular region A_i with the boundary contour γ_i. Let X_1 denote the convex core of R. This is constructed as follows. Let I_i be a component over $\gamma_i \subset \partial\mathbb{D}$; it is stabilized by an element $a_i \in G$. The axis C_i^* of a_i ends at the endpoints of I_i and lies over C_i. Cut out of $\mathbb{D}$ the region bounded by $C_i^* \bigcup I_i$. When this is done for all boundary contours and their lifts to $\mathbb{D}$, what is left is the lift of X_1.

The convex core X_1 is itself a compact bordered Riemann surface with the same genus, number of punctures, and number of boundary contours as R. Introduce on the interior of X_1 its complete hyperbolic metric. Repeat the process; that is, let X_2 be the convex core of X_1. There results a nested sequence of subsets of $\mathbb{D}$:

$$\mathbb{D} \supset X_1 \supset X_2 \supset \cdots.$$

Set $Z = \bigcap_{i=1}^{\infty} X_i$. Bers first raised the problem: Describe Z. Following the insight provided by the special cases in [Earle 1993], Jianguo Cao [1994] proved that Z has no interior, and that Z is the Hausdorff limit of souls $S(X_i)$.

Cao defines the *soul* $S(R)$ of R (or of any bordered surface) to be the set of points $z \in R$ such that there are at least two distinct shortest geodesic segments from z to $\bigcup C_i$. It contains $\bigcup C_i$. If there are no punctures, the soul is compact.

The soul is a union of geodesic arcs and is a deformation retract of R.

Explore this situation with the goal of gaining more precise information about Z, and finding a purely geometric proof of Cao's results. What about 3D?

3-34. *Extension from $\Omega(G)$ to $\mathbb{S}^2$.* Suppose as in Section 3.7.2 that F_2 is a quasiconformal map of $\mathbb{S}^2$, with $F_2 : \Omega(G) \to \Omega(H)$, that induces the isomorphism $\varphi : G \to H$ between geometrically finite groups. Suppose that $F : \Omega(G) \to \Omega(H)$ is quasiconformal, homotopic on $\Omega(G)$ to the restriction of F_2, and also induces φ. Using the density of the loxodromic fixed points in $\Lambda(G)$ show that

Lemma 3.14.5. *F has a continuous extension to a homeomorphism of $\mathbb{S}^2$ that satisfies $F(\zeta) = F_2(\zeta)$ for all $\zeta \in \Lambda(G)$.*

Set $H = {F_2}^{-1} \circ F : \Omega(G) \to \Omega(G)$. The map H is homotopic to the identity on each component of $\Omega(G)$, induces the identity automorphism of G, and is equal to the identity on $\Lambda(G)$. Let γ_z be a shortest geodesic from z to $H(z)$. There is a constant $C_1 < \infty$ such that $L_h(z) = d_h(z, H(z)) < C_1$ for all $z \in \Omega(G)$ (lift from the quotient). Here $d_h(\cdot, \cdot)$ denotes the shortest hyperbolic distance on $\Omega(G)$. From this it follows if $z \to \zeta \in \Lambda(G)$ in the spherical metric, then $\lim \gamma_z = \zeta$ uniformly on γ_z. That is, there exists a constant C_2 such that $d(\gamma_z, \Lambda(G)) < C_2$ for all $z \in \Omega(G)$, where $d(\cdot,\cdot)$ denotes spherical distance. Hence $d(w,\Lambda(G)) < C_2 d(z,\Lambda(G))$ for all $w \in \gamma_z$ and some $C_2 < \infty$. (Actually we only need these estimates for $z \in \Omega(G)$ near a point $\zeta \in \Lambda(G)$.)

I am grateful to Vlad Markovic for allowing inclusion of his unpublished result as follows.

Proposition 3.14.6 (Markovic). *H is quasiconformal on $\mathbb{S}^2$; hence F itself is the restriction to $\Omega(G)$ of an equivariant quasiconformal map of $\mathbb{S}^2$.*

Proof. Markovic's proof is as follows. Set $X = \Lambda(G)$. From [Pommerenke 1984] we know that X has the property of uniform perfectness, see Exercise 1-30. That is, for the hyperbolic metric $\rho(w)|dw|$ in each component of $\Omega(G)$ and some constant $C_3 > 0$,

$$\frac{C_3|dw|}{d(w,X)} < \rho(w)\,|dw| < \frac{2|dw|}{d(w,X)}.$$

Upon integrating over a shortest geodesic γ_z of hyperbolic length $L_h(z)$ from z to $H(z)$, we find that $C_3 d(z, H(z) \leq L_h(z)\sup_{w\in\gamma_z} d(w,X) < L_h C_2 d(z,X)$. In other terms, $d(z,H(z)) < C_4 d(z,\zeta)$ for any $\zeta \in X$. Now $d(H(z),\zeta) \leq d(H(z),z) + d(z,\zeta)$. Consequently for some constant C_5, $d(H(z),\zeta) < C_5 d(z,\zeta)$. The same holds if we replace z by $H(z)$. We conclude that

$$\frac{d(H(z),\zeta)}{C_5} \leq d(z,\zeta) \leq C_5 d(H(z),\zeta).$$

So the ratio of distances to ζ is uniformly bounded between 0 and ∞ as $z \to \zeta$. We are now in position to apply the geometric definition §2.8 of quasiconformality to show that $H(\zeta)$ is quasiconformal at ζ. Since ζ was arbitrarily chosen this proves H is quasiconformal on $\Lambda(G)$. □

3-35. *Intersection number estimates* [Fathi et al. 1979, pp. 58–59]. Let R_0 denote the result of removing the universal horodisks from R. Prove:

- There is a constant $C > 0$ such that for any two simple closed geodesics α, β their intersection number satisfies

$$\iota(\alpha,\beta) < C\,\mathrm{Len}(\alpha)\,\mathrm{Len}(\beta).$$

- Let $\{\tau_i\}$ be a finite system of simple closed geodesics and simple arcs that cut R_0 into simply connected regions. There exists a constant $c = c(\cup\tau_i)$ such that for any simple closed geodesic α in R_0,

$$\sum \iota(\alpha,\tau_i) \geq c\,\mathrm{Len}(\alpha).$$

Hints: For the first, cover α and β by ϵ-disks thereby dividing them into short geodesic segments. Each segment intersects another at most at one point. Show that

$$\iota(\alpha,\beta) < (\frac{\mathrm{Len}(\alpha)}{\epsilon}+1)(\frac{\mathrm{Len}(\beta)}{\epsilon}+1).$$

For the second, let $c = 1/L$ where L is the length of the longest simple arc in the simply connected regions.

3-36. *Interval exchange transformations* [Masur 1982; Bonahon 2001]. Let $I \subset \mathbb{R}$ be the interval (0,1]. Write $I^{\pm}$ for its upper and lower edges. Suppose $\{I_i^+\}$ is a partition of I^+ into n half-closed intervals $\{[a_{i-1}, a_i)\}$ of various lengths, with $a_0 = 0$, $a_n = 1$. Take the same sequence of intervals on I^- but then permute them in any way. Label the result by the notation $\{I_i^-\}$ where I_i^- has the same length as I_i^+. The corresponding *internal exchange transformation* J is the piecewise euclidean isometry that maps I_i^+ onto I_i^-, $1 \le i \le n$. The map is one-to-one, except two-to-one at the endpoints of the closures of the intervals. We must chose the permutation so the interval exchange does not reduce to an exchange of fewer intervals, that is, so that J is not continuous at any interval endpoint.

There is a naturally associated closed Riemann surface R: View the complement of I in $\mathbb{S}^2$ as a polygon with n-pairs of edges $I_i^{\pm}$. Identify each pair of edges by the direction preserving isometry; akin to what we did by "rolling up" fundamental regions by their edge identifications. The resulting surface will have singular points coming from the endpoints of the intervals. But there will be a natural complex structure at these points as well which maps the local neighborhoods into $\mathbb{C}$.

The vertical euclidean lines give rise to a *measured foliation* of R. Namely, except for a countable number of points, given $x \in I$ the forward and backward orbit $J^{\pm n}(x)$ will not hit an interval endpoint. These generic points will lie on a leaf of a foliation of R by vertical lines. The differential dx is the local vertical measure of the foliation.

The foliation is turned into a measured lamination by showing in the universal cover, the leaves have endpoints on $\partial\mathbb{H}^2$ and replacing each leaf by a geodesic.

By adjusting the interval lengths one can obtain minimal laminations, and uniquely ergodic ones as well. See [Masur 1982] for more details and further references.

3-37. *Horocyclic foliations.* Assume we have a closed surface S, a hyperbolic metric g on S, and a maximal geodesic lamination Λ such that all complementary regions are ideal triangles. From this data we will construct a measure foliation called a *horocyclic foliation*.

First foliate each ideal triangle as follows. Model the triangle by an ideal triangle in the disk model $\mathbb{D}$ whose sides have equal euclidean lengths. Foliate a neighborhood of each vertex v_i by a family of arcs contained in concentric circles with center at v_i. Do this symmetrically about all three vertices. We are left with a small central curved triangle that is not foliated.

Using our model example, foliate $S \setminus \Lambda$. The leaves joint together to form a family of mutually disjoint open arcs of infinite length in S, each orthogonal to the leaves of Λ. Collapse the finitely many central triangles to points. This results in a singular foliation of a new surface S_0 equivalent to S with singular points at the collapsed triangles. Replace S by S_0.

The hyperbolic metric g determines a transverse measure μ_g by measuring vertical distances along the leaves of Λ. Thurston [1998] proved that the map $g \mapsto (\Lambda_g, \mu_g)$ is a homeomorphism of Teich(S) onto its image in measure foliation space.

4

Algebraic and geometric convergence

The focus of this chapter is on sequences of kleinian groups, typically sequences that are becoming degenerate in some way. For these, it is necessary to carefully distinguish between convergence of groups and convergence of quotient manifolds. The former has to do with sequences of groups whose generators converge, the latter with sequences of groups whose fundamental polyhedra converge. Our work in this chapter will enable us to describe the set of volumes of finite volume hyperbolic 3-manifolds. In preparation for this discussion, we will introduce the operation called Dehn surgery.

4.1 Algebraic convergence

In this section we will prove the two theorems which provide the basis for working with sequences of groups.

Let Γ be an abstract group and $\{\varphi_n : \Gamma \to G_n\}$ be a sequence of *homomorphisms* (also called *representations*) $\{\varphi_n\}$ of Γ to groups G_n of Möbius transformations. Suppose for each $\gamma \in \Gamma$, $\lim_{n\to\infty} \varphi_n(\gamma) = \varphi(\gamma)$ exists as a Möbius transformation. Then the sequence $\{\varphi_n\}$ is said to *converge algebraically* and its *algebraic limit* is the group $G_\infty = \{\varphi(\gamma) : \gamma \in \Gamma\}$; $\varphi : \Gamma \to G_\infty$ is a homomorphism. When we say a sequence of groups converges algebraically, we are assuming that behind the statement is a sequence of homomorphisms generating the sequence.

In particular, a sequence of r-generator groups $G_n = \langle A_{1,n} A_{2,n} \dots A_{r,n}\rangle$ is said to converge *algebraically* if $A_k = \lim_{n\to\infty} A_{k,n}$ exists as a Möbius transformation, $1 \le k \le r$. Its *algebraic limit* is the group $G = \langle A_1, A_2, \dots A_r\rangle$. To make this terminology consistent with that used above, refer to the free group F_r on r-generators and express G_n as the sequence of representations $\varphi_n : F_r \to G_n$ determined by sending the k-th generator of F_r to $A_{k,n}$.

If the sequence $\{G_n\}$ consists of elementary groups, the limit may or may not be discrete. However in most interesting cases, the sequence consists of nonelementary groups. For nonelementary groups, the convergence is controlled as spelled out by the following fundamental results.

Theorem 4.1.1 [Jørgensen 1976; Jørgensen and Klein 1982]. *Let $\{G_n\}$ be a sequence of r-generator nonelementary kleinian groups converging algebraically to the group G. Then G is also a nonelementary kleinian group, and the map $A_k \to A_{k,n}$, $1 \le k \le r$, determines a homomorphism $\phi_n : G \to G_n$ for all large indices n.*

In general ϕ_n will not be an isomorphism. For example, a sequence of elliptic transformations $\{A_{k,n}\}$ may converge to a parabolic transformation A_k. In the opposite direction, Theorem 4.1.1 implies that if some A_k is elliptic of order r, then so is $A_{k,n}$ for all large n.

In contrast to Theorem 4.1.1, in applications we frequently work with isomorphisms from a fixed group:

Theorem 4.1.2 [Jørgensen 1976]. *Suppose G is a nonelementary kleinian group and $\{\theta_n : G \to G_n\}$ is a sequence of isomorphisms onto kleinian groups G_n. Assume that for each element $g \in G$, $\lim_{n\to\infty} \theta_n(g) = \theta(g)$ exists as a Möbius transformation. Then $G_\infty = \{\theta(g) : g \in G\}$ is a nonelementary kleinian group and $\theta : G \to G_\infty$ is an isomorphism.*

In Theorem 4.1.2, we do not need to require that G be finitely generated.

These two theorems are consequences of Jørgensen's inequality.

Proof of Theorem 4.1.2. First we will show $\theta : G \to \theta(G)$ is an isomorphism. If it is not, $\theta(g)$ is the identity for some $g \ne \mathrm{id} \in G$. Since θ_n is an isomorphism, if $g \in G$ has finite order, $\theta_n(g)$ and in the limit $\theta(g)$ must have exactly the same order. Therefore if $\theta(g) = \mathrm{id}$, g must have infinite order. Since G is not elementary, there is an element $h \in G$ also of infinite order but without a common fixed point with g. Each group $\langle \theta_n(g), \theta_n(h)\rangle$ is also nonelementary, since a nonelementary discrete group cannot be isomorphic to an elementary discrete group. But then we are in violation of Jørgensen's inequality (2.1) for all large n.

Next we will show that $\theta(G)$ is discrete. If not, there is a sequence $\{g_k \ne \mathrm{id} \in G\}$ such that $\lim_{k\to\infty} \theta(g_k) = \mathrm{id}$. There is a sequence $n = n(k)$ so that $\lim_{k\to\infty} \theta_n(g_k) = \mathrm{id}$. We may assume that either all g_k lie in the same cyclic subgroup or else their fixed points are mutually disjoint. In either case we can find an element $h \in G$ of infinite order whose fixed points are distinct from those of all but a finite number of the elements g_k. The sequence of nonelementary groups $\langle \theta_n(h), \theta_n(g_k)\rangle$ for large k is in violation of Jørgensen's inequality.

Finally we have to show that $\theta(G)$ is nonelementary. That is now easy because an elementary discrete group is a finite extension of an abelian group. □

Proof of Theorem 4.1.1. We start with a sequence of lemmas:

Lemma 4.1.3. *If each of the four Möbius transformations A, B, AB, $ABA^{-1}B^{-1}$ is elliptic or the identity, then they have a common fixed point in $\mathbb{H}^3$.*

Proof. If A and B commute, then according to Lemma 1.5.2, they either have the same axes, or each is of order two and their axes are orthogonal at a common point of intersection. In either case the conclusion is obvious. So assume they do not commute. Find U so that the conjugates UAU^{-1}, UBU^{-1} are such that the former has fixed points $0, \infty$. Then conjugate both by $V = \begin{pmatrix} \sqrt{c} & 0 \\ 0 & 1/\sqrt{c} \end{pmatrix}$, where c is the lower left entry of UBU^{-1}. Rename the results by A, B so as to end up with the following:

$$A = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}, \quad B = \begin{pmatrix} a & b \\ 1 & d \end{pmatrix}, \quad 2\theta \not\equiv 0,\ ad - b = 1.$$

Now $\mathrm{tr}(B) = a + d = r_1$ and $\mathrm{tr}(AB) = e^{i\theta}a + e^{-i\theta}d = r_2$ are real while $e^{2i\theta} \neq 1$. Solving the two equations for a and d we find that $a = \overline{d}$. Since B is elliptic, its trace satisfies $-2 < \mathrm{tr}(B) < 2$. Hence the fixed points of B, namely $\frac{1}{2}\big(a - d \pm \sqrt{\mathrm{tr}^2(B) - 4}\,\big)$, are purely imaginary. Further the product of the fixed points is $1 - ad = -b$.

Next we find for the commutator that $\mathrm{tr}(ABA^{-1}B^{-1}) - 2 = 4b\sin^2\theta$. Since this is elliptic as well, we must have $b < 0$. Since the product of the fixed points of B is positive, they lie on opposite sides of $z = 0$. That is, in the upper half-space model of $\mathbb{H}^3$, the axes of A and B intersect. The point of intersection is fixed by both, and by the group they generate. $\square$

Lemma 4.1.4. *Suppose that A, B are elliptic and that their axes intersect properly in $\mathbb{H}^3$. Then the plane P containing their axes does not contain the axis of AB.*

Proof. [Gallo et al. 2000, Lemma 3.4.3] Fix a point x on the axis of the elliptic AB which does not lie on the axis of B. Set $y = B(x)$ so that $A(y) = x$. Let P' denote the plane which is the perpendicular bisector of the segment $[x, y]$ so that x and y are equidistant from P' in the hyperbolic metric. But x and y are also equidistant from each point on the axis of B, since B is a rotation about its axis. All points equidistant from x and y lie in P' so that the axis of B lies in P'. But x and y are also equidistant from the axis of A so that lies in P' as well. Therefore $P' = P$. But x, which lies in the axis of AB, does not lie in P'. $\square$

We will digress from the proof of Theorem 4.1.1 to draw the following important corollary:

Corollary 4.1.5. *A group G, discrete or not, composed only of elliptic elements either has a common axis or it has a unique common fixed point in $\mathbb{H}^3$.*

Proof. Suppose $A, B \in G$ have distinct axes ℓ_A, ℓ_B. Lemma 4.1.3 shows that they intersect in a point $x \in \mathbb{H}^3$. Let P be the plane they span. Consider a third element $C \in G$ with axis ℓ_C distinct from ℓ_A, ℓ_B. If $x \notin \ell_C$, then $\ell_C \subset P$. So all other axes from G either pass through x or lie in P. On the other hand by Lemma 4.1.4, the axis ℓ_{AB} of AB, which intersects ℓ_A at x, does not lie in P. So the plane P' spanned by ℓ_A, ℓ_{AB} does not coincide with P. But repetition of our argument shows that ℓ_C lies in P' as well, a contradiction. We have shown that either all elements of G have the same axis, or the set of axes of elements of G have a single common point of intersection in $\mathbb{H}^3$. $\square$

Lemma 4.1.6. *If $\langle A, B\rangle$ is nonelementary, then at most one of the three elements A, B, AB is elliptic of order two.*

Proof. An element of order two is conjugate to $z \mapsto -z$ whose axis in the upper half-space model is the half-line rising from $z = 0$. If both A and B are elliptic of order two and their axes do not coincide, there is a unique common perpendicular line ℓ to the two axes. Since A and B are rotations by π about their axes, each maps ℓ onto itself by rotating it by π about the crossing point with its axis. The cyclic subgroup $\langle AB\rangle$ which maps ℓ onto itself without reversing direction has index two. The bottom line is that the group $\langle A, B\rangle$ is elementary, a contradiction. □

Lemma 4.1.7. *Suppose $A_1, \ldots, A_r$ generate an infinite, discrete group G. The set*

$$\{A_i,\ A_iA_j,\ A_iA_jA_k,\ (A_iA_j)A_k(A_iA_j)^{-1}A_k^{-1}:\ i, j, k = 1, \ldots, r\}$$

contains an element of infinite order.

Proof. The assertion is true for $r = 1$. For $r \geq 2$, assume the assertion is false so that all the listed elements have finite order. We will show that this implies they have a common fixed point in $\mathbb{H}^3$. This in turn will imply that G is a finite group, in contradiction to the hypothesis.

Use our current hypothesis and apply Lemma 4.1.3 to $A = A_iA_j$, $B = A_i$. We see that A_iA_j, A_i have a common fixed point x in $\mathbb{H}^3$ and therefore A_j fixes x as well. Consequently the axes of the generators $\{A_i\}$ pairwise intersect.

Choose a point $x \in \mathbb{H}^3$ at which a maximal number of generator axes intersect, say the axes of $A_1, \ldots, A_m$, $2 \leq m \leq r$. If $m = r$ we are finished. If $m < r$, then the axis ℓ of A_{m+1} does not pass through the common point x of its predecessors. We may assume that the axes of A_1 and A_2 meet ℓ at different points. Let P be the plane spanned by the axes of A_1, A_2; necessarily P contains x and ℓ. According to Lemma 4.1.4, the axis of A_1A_2, which goes through x, does not lie in P. So the axis of A_1A_2 is disjoint from ℓ, in contradiction to Lemma 4.1.3. □

Incidentally we have confirmed the following:

Corollary 4.1.8. *A discrete group in which every element is elliptic is a finite group.*

Lemma 4.1.9. *Suppose that $A_1, \ldots, A_r$ generate a nonelementary, discrete group G and that $g \in G$ is loxodromic or parabolic. Then $H_i = \langle A_i, g\rangle$ is nonelementary for at least one index i.*

Proof. Suppose to the contrary that for each index, H_i is elementary. If g is loxodromic, each A_i must fix or interchange the fixed points of g. Thus G itself must be elementary since it fixes the set of two fixed points of g. If g is parabolic, each A_i fixes the fixed point of g and again G is elementary. □

Lemma 4.1.10. *Suppose that $A = \lim A_n$, $B = \lim B_n$ for two sequences of Möbius transformations where each group $G_n = \langle A_n, B_n\rangle$ is discrete. Then:*

(*a*) *If G_n is nonelementary for all indices, $A \neq$ id.*

(*b*) *If G_n is nonelementary for all indices and A is elliptic, its order is finite.*
(*c*) *If neither A nor B has order two, then A and B have a common fixed point on $\mathbb{S}^2$ if and only if A_n and B_n also do so for all large indices.*

Proof. We will first prove part (c). The hypothesis implies that for all large n, A_n, B_n are not elliptic of order two. First assume that A_n and B_n have a common fixed point on $\mathbb{S}^2$ for all large indices. The trace of their commutator is $+2$ by Lemma 1.5.1. Thus by continuity the commutator of A and B also has trace $+2$ and therefore A and B have a common fixed point as well.

Conversely, suppose A and B have a common fixed point so that the trace of their commutator $K = ABA^{-1}B^{-1}$ is $+2$. Also $K = \lim K_n = \lim A_n B_n A_n^{-1} B_n^{-1}$. For all large indices, either A_n and B_n have a common fixed point or K_n is not parabolic, since the trace $\neq -2$. If the former case occurs we are finished. So assume that for all large n, K_n is not parabolic.

Since A, B, K all share a fixed point,

$$\mathrm{tr}^2(K) - 4 = \mathrm{tr}(KAK^{-1}A^{-1}) - 2 = \mathrm{tr}(KBK^{-1}B^{-1}) - 2 = 0.$$

For all large indices then,

$$\left|\mathrm{tr}^2(K_n) - 4\right| + \left|\mathrm{tr}(K_n A_n K_n^{-1} A_n^{-1}) - 2\right| < 1,$$
$$\left|\mathrm{tr}^2(K_n) - 4\right| + \left|\mathrm{tr}(K_n B_n K_n^{-1} B_n^{-1}) - 2\right| < 1.$$

Since K_n is not parabolic and A_n and B_n are not of order two, according to Theorem 2.1(i), G_n is cyclic or a finite abelian extension of a cyclic group. Now K_n has two fixed points and neither A_n nor B_n exchange them. The remaining possibility is that A_n and B_n share the fixed points of K_n. This completes the proof of (c).

Part (a) is a direct consequence of Jørgensen's inequality. To prove (b), assume A is elliptic. If A is elliptic of infinite order then for some q, A^q is close to id. If $A_n^q \neq \mathrm{id}$, it is elliptic or loxodromic. For some q and all large indices,

$$\left|\mathrm{tr}^2(A_n^q) - 4\right| + \left|\mathrm{tr}(A_n^q B_n A_n^{-q} B_n^{-1}) - 2\right| < 1.$$

Either part (i) or (ii) of Theorem 2.1 applies. If it is (i), then G_n is elementary. If it is (ii), then B_n is elliptic of order two and interchanges the fixed points of A_n^q, and hence of A_n. Therefore $\langle A_n, B_n\rangle$ is elementary, a contradiction. □

We can now continue with the proof of Theorem 4.1.1. The hardest part is to prove that G is not elementary. The case $r \geq 3$ can be reduced to the case $r = 2$. For if $r \geq 3$, according to Lemma 4.1.7, there exists a loxodromic or parabolic $g_n \in G_n$ whose length as a word in the generators $\{A_{i,n}\}$, $1 \leq i \leq r$, is uniformly bounded in n. Applying Lemma 4.1.9, for each n there is a nonelementary subgroup H_n of G_n generated by g_n and some generator $A_{i,n}$. Since the word length of g_n is uniformly bounded terms of the given generators, a subsequence of the two generator groups H_n converges algebraically to a subgroup H of G. If H is nonelementary, so is G.

So we may assume that $r = 2$ and $G_n = \langle A_{1,n}, A_{2,n}\rangle$ converges algebraically to $G = \langle A_1, A_2\rangle$. In view of Lemma 4.1.6 after rearranging some more if necessary, we may assume that for all n, neither $A_{1,n}$ nor $A_{2,n}$ has order two. Then neither A_1 or A_2 can have order two, for if say $A_1^2 = \text{id}$, replace G_n by $\langle A_{1,n}^2, A_{2,n}\rangle$ to get a contradiction to Lemma 4.1.10(a).

Since G_n is nonelementary, $A_{1,n}$ and $A_{2,n}$ have distinct fixed points. By Lemma 4.1.10(c), A_1 and A_2 have distinct fixed points as well.

Among the four elements A_1, A_2, A_1A_2, $A_1A_2A_1^{-1}A_2^{-1}$ of G, there is an element X of infinite order. Otherwise these elements would be elliptic of finite order, and the same would hold for the corresponding elements of G_n for large indices. But by Lemma 4.1.3, G_n would be a finite group. By Lemma 4.1.10(b), X cannot be elliptic.

If X is parabolic, then at least one of the parabolic elements $A_1XA_1^{-1}$, $A_2XA_2^{-1}$ has no common fixed point with X. For otherwise A_1 and A_2 have a common fixed point. This makes G nonelementary.

If X is loxodromic, we claim that there exists $Y \in G$ such that X and YXY^{-1} have no fixed points in common.

To establish this claim, we will first investigate what happens if for $Y \in G$, X and YXY^{-1} do have a common fixed point. Applying again Lemma 4.1.10(c), we see that the corresponding elements in G_n also have a common fixed point, for all large indices. These approximants are loxodromic. Since G_n is discrete, two loxodromic elements cannot have exactly one fixed point in common by Lemma 2.3.1(ii),(iii). Therefore X and YXY^{-1} have both fixed points in common. Unless Y has order two and interchanges the fixed points of X, X and Y have the same fixed points too. Consequently by choosing Y as either A_1 or A_2 we obtain the desired result that X and YXY^{-1} have no fixed points in common. We conclude that G is not elementary.

Now return to the hypothesis of Theorem 4.1.1. We are given a sequence $G_n = \langle A_{1,n}, \dots, A_{r,n}\rangle$ such that $\lim_{n\to\infty} A_{k,n} = A_k$ with $G = \langle A_1, \dots, A_r\rangle$. We will use the correspondence $\phi : G \to G_n$ generated by $\phi_n : A_k \to A_{k,n}$.

We are ready to prove that G is discrete. Suppose otherwise. Then there exists a sequence of elements $B_k \in G$ with $\lim B_k = \text{id}$. We may assume that no B_n has order two. Since G is nonelementary, according to Exercise 2-1 there are two loxodromic elements $g_1, g_2 \in G$ without a common fixed point. By Lemma 4.1.10(c), g_i and B_k have a common fixed point if and only if $\phi_n(g_i)$ and $\phi_n(B_k)$ also do, for large n. Since G_n is discrete, this can occur only if the elements have the same fixed points, for neither is of order two. In this case g_i and B_k also have the same fixed points. The bottom line is that we can pick an infinite subsequence so that g_1, say, has no fixed point in common with $\{B_m\}$. Likewise $\phi_n(g_1)$ and $\phi_n(B_m)$ have no fixed points in common so generate a nonelementary subgroup. Now $\{\langle\phi_n(g_1), \phi_n(B_m)\rangle\}$ violates Jørgensen's inequality. Hence G is discrete.

The last step is to show that the correspondence $\phi_n : A_k \to A_{k,n}$ can be extended to a homomorphism $\phi_n : G = \langle A_1, \dots, A_r\rangle \to G_n$. A necessary and sufficient condition that extension of ϕ_n to a homomorphism $G \to G_n$ is possible is that for each "relation" $R = \prod_k A_{i_k}^{m_k} = \text{id}$ in G also $\phi_n(R) = \prod_k A_{i_k,n}^{m_k} = \text{id}$. First of all, by Theorem 2.5.3, there are only a finite number of relations in G: more precisely, there are a finite number of relations in G such that every other relation is a consequence of those. For if to the contrary $\phi_n(R) \neq \mathrm{id}$, application of Lemma 4.1.10(a) results in a contradiction. $\square$

There is another interesting corollary. Denote the space of ordered r-tuples of Möbius transformations which generate nonelementary groups by V_r. Let D_r be the subset consisting of discrete groups.

Corollary 4.1.11 [Jørgensen 1976]. *Each component of D_r consists of mutually isomorphic groups.*

Proof. Choose a component D and a group $G \in D$. Let X denote the set of all homomorphic images ϕ of G in D. By Theorem 4.1.1, X is relatively open in D. It is also closed in D. Therefore $X = D$. The same argument holds upon replacing G by any $G_1 \in D$. We conclude that G and G_1 are isomorphic. See also Section 5.1. $\square$

4.2 Geometric convergence

Algebraic convergence deals not with geometry but with convergence of group generators. It is possible that in a sequence of groups $\{G_n\}$ there are words $W_n \in G_n$ in the generators, whose length increases without bound as $n \to \infty$, yet which converge to a Möbius transformation. Such phenomena are not detected by focusing on convergence of generators. Instead the phenomenon impacts the behavior of the sequence of quotient manifolds. From the point of view of a manifold, the generators of the fundamental group are rather arbitrarily chosen loops. What is fundamental are the geometrical quantities that determine its "shape". If we have a sequence of manifolds, we need a framework for discussing convergence to a limiting manifold.

If $\{G_n\}$ is a sequence of groups of Möbius transformations, define its *envelope* as

$$\mathrm{Env}\{G_n\} = \{g \in \mathrm{PSL}(2, \mathbb{C}) : g = \lim g_n, \ g_n \in G_n\}.$$

It follows that $\mathrm{Env}\{G_n\}$ is itself a group.

Lemma 4.2.1. *If each G_n is discrete, then either $H = \mathrm{Env}\{G_n\}$ is elementary, or it is a nonelementary, discrete group.*

Proof. According to Corollary 2.2.1 a group is discrete if and only if every two generator subgroup is discrete. Assume that H is not elementary. Then given an element h_1 of infinite order there is another h_2 without a common fixed point. If the nonelementary subgroup $\langle h_1, h_2\rangle$ were not discrete, we could find as in the final part of the proof of Theorem 4.1.1 that $h_1', h_2' \in \langle h_1, h_2\rangle$ with h_2' nearly the identity such that $\langle h_1', h_2'\rangle$ is nonelementary yet in violation of Jørgensen's inequality (2.1). Now h_1', h_2' are each limits of elements in G_n. For large n, the pair of approximants in G_n will generate a nonelementary subgroup of G_n yet violate Jørgensen's inequality, a contradiction. This proves that any two generator nonelementary subgroup of H is discrete. $\square$

We say that the sequence of groups $\{G_n\}$ *converges geometrically* (to $\mathrm{Env}\{G_n\}$) if and only if for every subsequence $\{G_{n_j}\}$ of $\{G_n\}$, $\mathrm{Env}\{G_{n_j}\} = \mathrm{Env}\{G_n\}$. In other words, $\{G_n\}$ converges geometrically to H if and only if (i) each $h \in H$ is the limit $h = \lim g_n,\ g_n \in G_n$, and (ii) whenever $\lim g_{n_j} = g$ exists for a subsequence $\{n_j\}$ then $g \in H$. Necessarily $H = \mathrm{Env}\{G_n\}$.

To justify use of the term "geometric convergence", and to give a precise meaning to the expression "convergent sequence of hyperbolic manifolds" we introduce the auxiliary concept of polyhedral convergence.

4.3 Polyhedral convergence

The sequence of discrete groups $\{G_n\}$ converges *polyhedrally* to the group H if H is discrete and for some point $\mathcal{O} \in \mathbb{H}^3$, the sequence of Dirichlet fundamental polyhedra $\{\mathcal{P}(G_n)\}$ centered at $\mathcal{O}$ converge to $\mathcal{P}(H)$ for H, also centered at $\mathcal{O}$, uniformly on compact subsets of $\mathbb{H}^3$.

We need to be more precise about the criterion for polyhedral convergence. Given $r > 0$, set

$$B_r \;=\; \{\vec{x} \in \mathbb{H}^3 : d(\mathcal{O}, \vec{x}) < r\},$$

where $d(\cdot,\cdot)$ denotes hyperbolic distance. We will work with the truncated polyhedra $\mathcal{P}_{n,r} = \mathcal{P}(G_n) \cap B_r$ and $\mathcal{P}_r = \mathcal{P}(H) \cap B_r$. A truncated polyhedron $\mathcal{P}_r$ has the property that its faces, that is the intersection with B_r of the faces of $\mathcal{P}$, are arranged in pairs, paired by the corresponding face pairing transformations of $\mathcal{P}$. Thus the projection of $\mathcal{P}_r$ into the quotient 3-manifold is a relatively compact submanifold, bounded by the projection of $\mathcal{P} \cap \partial B_r$ (Proposition 3.5.1).

The criterion for polyhedral convergence is as follows. Given any r sufficiently large, there exists $N = N(r) > 0$ such that (i) to each face pairing transformation h of $\mathcal{P}_r$, there corresponds a face pairing transformation g_n of $\mathcal{P}_{n,r}$ for all $n \geq N$ such that $\lim_{n\to\infty} g_n = h$, and (ii) if g_n is a face pairing transformation of $\mathcal{P}_{n,r}$ then the limit h of any convergent subsequence of $\{g_n\}$ is a face, edge or vertex pairing transformation of $\mathcal{P}_r$; in particular $h \neq \mathrm{id}$. In short, each pair of faces of $\mathcal{P}_r$ is the limit of a pair of faces of $\{\mathcal{P}_{n,r}\}$, and each convergent subsequence of a sequence of face pairs of $\{\mathcal{P}_{n,r}\}$, converges to a pair of faces, edges, or vertices of $\mathcal{P}_r$. We remark that it is possible that $\mathcal{P}_{n,r} = B_r$ for all large n. In this case the sequence of polyhedra converges to $\mathbb{H}^3$ itself.

If a given sequence of discrete groups is to converge polyhedrally, one must be allowed to conjugate the groups if necessary to find a point $\mathcal{O} \in \mathbb{H}^3$ that can effectively serve as center for all the polyhedra. We should be aware of the fact that a group can be conjugated so that for fixed $\mathcal{O}$, $\mathcal{P}_{\mathcal{O}}$ collapses. Namely conjugating $(az+b)/(cz+d)$ by $z \mapsto kz$ results in $k^{-1}(kaz+b)/(kcz+d)$. Its limit as $k \to \infty$ is 0, if $c \neq 0$.

The criterion needed is that there be a small ball about $\mathcal{O}$ that lies in the interior of the polyhedron for every group in the sequence. This is described in the following lemma.

Lemma 4.3.1. *To any infinite sequence of discrete groups $\{G_n\}$ corresponds a sequence of conjugates $\{A_n G_n {A_n}^{-1}\}$ which contains a polyhedrally convergent subsequence.*

Proof. Given $\mathcal{O} \in \mathbb{H}^3$, for each n choose a Möbius transformation A_n such that $G'_n = A_n G_n A_n^{-1}$ has the following property: Each truncated polyhedron $\mathcal{P}'_{n,r} = \mathcal{P}(G'_n)_r$ centered at $\mathcal{O}$ contains the ball B_δ centered at $\mathcal{O}$ for $r > \delta$. Here $\delta > 0$ is a fixed number given by the universal ball property (Proposition 3.3.4). Thus the sequence of polyhedra centered at $\mathcal{O}$ of the conjugate groups cannot collapse to a convex object without interior… which is certainly possible in general.

We claim that for fixed $r > \delta$ the number of faces of the truncated polyhedra $\{\mathcal{P}'_{n,r}\}$ is uniformly bounded as $n \to \infty$. The reason for this is that there is an upper bound on the number of mutually disjoint balls of hyperbolic radius δ that fit inside B_{3r}. Therefore there is an upper bound, independent of n, on the number of points in the orbit $G'_n(\mathcal{O})$ that lie in B_{3r}. A face pairing transformation of $\mathcal{P}'_{n,r}$ satisfies $d(\mathcal{O}, g_n(\mathcal{O})) < 2r$, and the segment $[\mathcal{O}, g_n(\mathcal{O})]$ pierces a face. Hence there is also a uniform bound M independent of n on the number of faces.

Consequently given $s > 0$, there exists a large $r = r(s)$ such that the orbit of $\mathcal{P}'_{n,r}$ under its face pairing transformations covers the ball B_s for all n. This is because there is a uniform bound on the length of words W in the face pairing transformations of $\mathcal{P}(G'_n)$, and the length of their segments $[\mathcal{O}, W(\mathcal{O})]$, required for the images of $\mathcal{P}(G'_n)$ to cover B_s. For sufficiently large r all of the elements W are also words in the face pairing transformations of the truncated polyhedra. The number of polyhedra meeting B_s is uniformly bounded in n by some $N < \infty$.

For fixed r and each n make a list of the face pairing transformation $\{g_{i,n}\}$ of $\mathcal{P}'_{n,r}$, $1 \le i \le M$ (by repetition we may assume there are M faces for each n). Take a subsequence of $\{n\}$ and relabel so that for each i, $h_i = \lim_{n\to\infty} g_{i,n}$ exists; $h_i \neq \mathrm{id}$ because $d(\mathcal{O}, g_{i,n}(\mathcal{O})) > 2\delta$. Correspondingly construct the polyhedron

$$\mathcal{P}_r^* = \{\vec{x} \in \mathbb{H}^3 : d(\mathcal{O}, \vec{x}) \le d(\vec{x}, h_i(\mathcal{O})),\ 1 \le i \le M\}.$$

Thus $B_\delta \subset \mathcal{P}_r^* \cap B_r = \lim \mathcal{P}'_{n,r}$.

Now take a sequence $r = r_k \to \infty$ and repeat the process for each r_k. We get a nested sequence of polyhedra $B_\delta \subset \mathcal{P}^*_{r_1} \subset \mathcal{P}^*_{r_2} \subset \cdots$. Set $\mathcal{P}_\infty = \bigcup_{i=1}^\infty \mathcal{P}^*_{r_i}$. The successive sets of side pairing transformations of the $\mathcal{P}^*_{r_k}$ are nested as well. Let $\{h_i\}$ denote the union. Let H denote the group they generate.

We claim that H is discrete, and $\mathcal{P}_\infty = \mathcal{P}_\mathcal{O}(H)$. Possibly $H = \{\mathrm{id}\}$ and $\mathcal{P}_\infty = \mathbb{H}^3$, the case that the groups $\{{G_n}'\}$ blow up completely.

First we claim that the orbit of $\mathcal{P}_\infty$ under H covers $\mathbb{H}^3$. As we have seen, given $s > 0$ there exists $r = r(s)$ such that the G'_n-orbit of $\mathcal{P}'_{n,r}$ covers the ball B_s for all n. For each n we can make a list of N transformations $W_{1,n}, W_{2,n}, \ldots, W_{N,n}$ such that $\bigcup_{i=1}^N W_{i,n}(\mathcal{P}'_{n,r}) \supset B_s$. Each $W_{i,n}$ is a word in the face pairing transformations of $\mathcal{P}'_{n,r}$, and the lengths are uniformly bounded as a function of s. Passing to a subsequence if necessary, each $W_i = \lim_{n\to\infty} W_{i,n}$ exists; necessarily $W_i \in H$. Therefore $(\cup_i W_i(\mathcal{P}_\infty)) \cap B_r$ covers B_s. Since s is arbitrarily chosen, our claim is established.

Next we claim that no two points in the interior of $\mathcal{P}_\infty$ are equivalent under H. For suppose that $W(x) = y$ for $x, y \in \mathrm{Int}(\mathcal{P}_\infty)$ and $W \in H$. The element W is a word in the generators $\{h_i\}$. For each n, let W_n denote the corresponding word in the approximants $\{g_{i,n}\}$ so that $W_n \in G'_n$ and $\lim W_n = W$, $\lim W_n(x) = y$. Choose $r > \max(d(\mathcal{O}, x), d(\mathcal{O}, y))$. Then for all large n, x and $W_n(x)$ lie in $\mathrm{Int}(\mathcal{P}'_{n,r})$. This is impossible unless $W_n = W = \mathrm{id}$.

We conclude that $\mathcal{P}_\infty$ is a fundamental polyhedron for H, and that H in turn is necessarily discrete. □

We can now justify our use of the term "geometric convergence". But first note that it is possible for a sequence of nonelementary discrete groups to converge geometrically and polyhedrally to an elementary group: here are two examples:

$$\left\{\left\langle z \mapsto -\frac{3z-1}{z-3}, z \mapsto z+n\right\rangle\right\}, \qquad \left\{\left\langle z \mapsto -\frac{1}{n^2}\frac{3n^2z-1}{n^2z-3}, z \mapsto z+\frac{1}{n}\right\rangle\right\}.$$

This is why in the following fundamental result we have to explicitly assume that the groups are nonelementary.

Proposition 4.3.2. *A sequence $\{G_n\}$ of kleinian groups converges geometrically to a nonelementary kleinian group if and only if it converges polyhedrally to a nonelementary kleinian group. The geometric and polyhedral limits are the same.*

Proof. Suppose first the sequence converges polyhedrally to H. If $h \in H$ then h is a word W in the face pairing transformations of $\mathcal{P}(H) = \mathcal{P}_\infty$. As in the proof of Lemma 4.3.1, the word is the limit of a sequence of words $W_n \in G_n$. Next we have to show that if for a subsequence $h = \lim g_k$, $g_k \in G_k$, then $h \in H$. Again we return to the proof of Lemma 4.3.1. Let $s = 2d(\mathcal{O}, h(\mathcal{O}))$, where the polyhedra are centered at $\mathcal{O}$. We showed that $B_s \subset \bigcup_{i=1}^N W_i(\mathcal{P}_\infty \cap B_r)$ where $W_i = \lim_{k\to\infty} W_{i,k}$ and $W_{i,k} \in G_k$. Therefore $h = W_i$, for some i, and is the limit as $k \to \infty$ of the corresponding word $W_{i,k}$. We conclude that H is the geometric limit.

Conversely, suppose H is the geometric limit of $\{G_n\}$. By the universal ball property of Proposition 3.3.4, there exists $\mathcal{O} \in \mathbb{H}^3$ such that the $h(B_\delta) \cap B_\delta = \varnothing$ for all $h \neq \mathrm{id} \in H$. We claim that the same property holds for G_n for all large n. Otherwise there would be a sequence $g_n \in G_n$ such that $g_n(B_\delta) \cap B_\delta \neq \varnothing$. A subsequence of $\{g_n\}$ converges to a Möbius transformation g_∞. If $g_\infty \neq \mathrm{id}$ then it would have to lie in H, a contradiction. If $g_\infty = \mathrm{id}$ we will find a contradiction to Jørgensen's inequality (2.1). Here we have to use the assumption that H is nonelementary. We can find two loxodromic transformations $h_1, h_2 \in H$ which have mutually disjoint fixed points (Exercise 2-1). Each is the limit $h_i = \lim h_{i,n}$, $h_{i,n} \in G_n$. For large n at least one of the $h_{i,n}$, say $h_{1,n}$, does not share a fixed point with g_n. Now we can apply Jørgensen's inequality to $\langle g_n, h_{1,n}\rangle$. The conclusion is that $\{G_n\}$ converges polyhedrally. □

We remark that the argument also applies in the following elementary situation. A sequence of cyclic loxodromic groups converges polyhedrally to a discrete parabolic group P if and only if it converges geometrically to P and no sequence of distinct elements converges to the identity. Here P may be of rank one or rank two. See Section 4.10.

Since geometric convergence makes no reference to a choice of center $\mathcal{O}$ for polyhedra, we can now remove any sign of dependence of polyhedral convergence on the choice of center $\mathcal{O}$.

Corollary 4.3.3. *If $\{G_n\}$ converges polyhedrally to the kleinian group H with one choice of center $\mathcal{O}$ for the polyhedra, it converges polyhedrally to H for any choice of center* (*which is not an elliptic fixed point of H*).

4.4 The geometric limit

We will need two lemmas. The first is a corollary of Theorem 4.1.1.

Lemma 4.4.1. *Suppose that $\{G_n\}$ is a sequence of nonelementary kleinian groups converging algebraically to G. There is no sequence of elements $g_k \in G_k$, $g_k \neq \mathrm{id}$, with $\lim g_k = \mathrm{id}$ or with $\lim g_k = g$ with g elliptic of infinite order.*

Proof. Present $G_n = \langle A_{1,n}, A_{2,n}, \ldots\rangle$, where $\lim A_{k,n} = A_k$ and no two generators have the same set of fixed points.

Case 1. g_n is elliptic for all large indices. For all large n, no generator $A_{k,n}$ can share exactly one fixed point with g_n. Otherwise $A_{k,n}$ would have to be parabolic and the order of g_n could not exceed six. Nor is it possible that every generator $A_{k,n}$ shares its fixed points with g_n or is of order two and interchanges the fixed points of g_n. For then G_n would be elementary. The conclusion is that for some k, $\langle g_n, A_{k,n}\rangle$ is nonelementary for all large indices, leading to a violation of Theorem 4.1.1.

Case 2. g_n is parabolic for all large indices. At most a finite number of elliptics can share its fixed point and at least one generator, say A_1 does not. This again leads to a violation of Jørgensen's inequality.

Case 3. g_n is loxodromic for all large indices. At most a bounded number of elliptics share a fixed point or interchange its fixed points, and at least one generator does neither. One is led to the usual contradiction. □

Lemma 4.4.2. *Suppose the sequence of nonelementary kleinian groups G_n converges algebraically to G. There exists a point $\mathcal{O} \in \mathbb{H}^3$ and $\varepsilon > 0$ such that, for a subsequence G_k, no element of G_k has a fixed point in the ball $B_\varepsilon(\mathcal{O})$ of radius ε about $\mathcal{O}$.*

Furthermore, there exists $\delta < \varepsilon$ such that for all large indices, $T_k(B_\delta(\mathcal{O}))$ is disjoint from $B_\delta(\mathcal{O})$, for all $T_k \neq \mathrm{id} \in G_k$.

Proof. We begin by showing that given $x \in \mathbb{H}^3$, there exists ε with the following property. There exists a point $x_n \in B_\varepsilon(x)$ such that any axis of G_n which intersects $B_\varepsilon(x)$ passes through x_n.

If this assertion were false there would be a sequence $\varepsilon_n \to 0$ and rotation axes ℓ_n, ℓ'_n of E_n, $E'_n \in G_n$ which intersect $B_{\varepsilon_n}(x)$ but don't have common point of intersection in $B_\varepsilon(x)$. We may assume there is convergence $E_n \to E,\ E'_n \to E'$ where E, E' both fix x. Also, their rotation axes converge to lines ℓ, ℓ' through x. By Lemma 4.4.1 E, E' are elliptic of finite order with rotation axes ℓ, ℓ'.

According to Theorem 4.1.1, or the Universal Elementary Property, $\langle E_n, E'_n\rangle$ is necessarily elementary for all large indices.

Now E_n, E'_n cannot share a fixed point on $\mathbb{S}^2$, for their commutator would then be parabolic and by Lemma 4.4.1 would remain parabolic in the limit. Yet the limit would have to fix x.

Nor can E_n, E'_n both be elliptic of order two with disjoint axes, for $E_n E'_n$ would then be loxodromic and the limit, which fixes x, could only be the identity, which is impossible again by Lemma 4.4.1.

The remaining alternative is that for all large indices, $\langle E_n, E'_n\rangle$ is a finite, noncyclic group with a common fixed point $x_n \in \mathbb{H}^3$. As a noncyclic finite group, the number of elements (the order) of $\langle E_n, E'_n\rangle$ is uniformly bounded. Therefore it is isomorphic to $\langle E, E'\rangle$, and $x_n \to x$.

From our argument we conclude that there exists $\varepsilon > 0$ and $x_n \in B_\varepsilon(x)$ such that any rotation axis of G_n that intersects $B_\varepsilon(x)$ passes through x_n, for all large indices. Moreover, the finite subgroups $\mathrm{Stab}(x_n) \subset G_n$ are isomorphic to the limit group denoted by $\mathrm{Stab}(x)$ and $\lim x_n = x$.

There are only a finite number of possibilities for $\mathrm{Stab}(x)$, unless it is cyclic or a $\mathbb{Z}_2$ extension of a cyclic group. Find $\mathcal{O} \in B_\varepsilon(x)$ and $\varepsilon_1 < \varepsilon$ such that $T B_{\varepsilon_1}(\mathcal{O}) \cap B_{\varepsilon_1}(\mathcal{O}) = \varnothing$ for all $T \neq \mathrm{id} \in \mathrm{Stab}(x)$. This property will persist for $\mathrm{Stab}(x_n)$, all large n.

Now consider the second assertion of Lemma 4.4.2. If it were false, corresponding to a sequence $\delta_n \to 0$ there would be a sequence $T_k \neq \mathrm{id} \in G_k$, $k = k(n)$, with

$$T_k(B_{\delta_n}(\mathcal{O})) \cap B_{\delta_n}(\mathcal{O}) \neq \varnothing.$$

Take a convergent subsequence, again labeled $\{T_k\}$. Its limit $T = \lim T_k$ fixes $\mathcal{O}$ but its approximates have no fixed point in $B_{\delta_n}(\mathcal{O})$. Therefore $T = \mathrm{id}$, again a violation of Lemma 4.4.1. □

Theorem 4.4.3 [Jørgensen and Marden 1990]. *Suppose the nonelementary kleinian groups G_n converge algebraically to G. Then there is a geometrically convergent subsequence $\{G_k\}$. The limit H of any geometrically convergent subsequence contains G; consequently $\mathcal{M}(G)$ is a covering manifold of $\mathcal{M}(H)$.*

If the geometric limit H is finitely generated, there is a sequence of homomorphisms to its approximants $\psi_k : H \to G_k$, for all large k, such that $\lim \psi_k(h) = h$ for all $h \in H$. In addition if G is finitely generated, then $\psi_k(H) = G_k$.

Proof. Set $G_n = \langle g_{1,n}, g_{2,n}, \ldots\rangle$ and $G = \langle g_1, g_2, \ldots\rangle$ with $g_i = \lim_{n\to\infty} g_{i,n}$. According to Lemma 4.4.1 there is no subsequence $h_k \in G_k$, $h_k \neq \mathrm{id}$, with $\lim h_k = \mathrm{id}$. Thus if the groups G_n contain no elliptic elements, about any given point $\mathcal{O} \in \mathbb{H}^3$, there is a small ball B_ε which is contained in every polyhedron $\mathcal{P}(G_n)$ centered at $\mathcal{O}$. In this case we can find a polyhedrally convergent subsequence as in Lemma 4.3.1.

When the groups G_n contain elliptic elements, Lemma 4.4.2 tells us that the ball $B_\varepsilon(\mathcal{O})$, for some $\mathcal{O} \in \mathbb{H}^3$, is such that for a subsequence, no element of G_k has a fixed point in $B_\varepsilon(\mathcal{O})$. Then Lemma 4.4.2 tells us more strongly that for some $\delta < \varepsilon$, $T B_\delta(\mathcal{O}) \cap B_\delta(\mathcal{O}) = \varnothing$, for all $T \neq \mathrm{id} \in G_k$. So $B_\delta(\mathcal{O})$ will lie in the Dirichlet region for each G_k centered at $\mathcal{O}$.

Thus in all cases there is a subsequence $\{G_k\}$ that converges polyhedrally to a group H.

Given a compact subset $X \subset \mathbb{H}^3$, we claim that there exists $r > 0$ and N with the following property: X is covered by the images of the truncated polyhedron $\mathcal{P}(G_k)_r$ under all words of length $\leq N$ in the face pairing transformations of $\mathcal{P}(G_k)_r$, for all large k.

To see why, choose a larger compact set $X' \supset X$ containing X in its interior. For large enough r, N, the orbit $\mathcal{Q}_N$ of $\mathcal{P}(H)_r$ under words of length $\leq N$ in the face pairing transformations of $\mathcal{P}(H)_r$ covers X'. When k is large, $\mathcal{P}(G_k)_r$ is close to $\mathcal{P}(H)_r$ since the faces of $\mathcal{P}(G_k)_r$ converge to those of $\mathcal{P}(H)_r$. The corresponding orbit $\mathcal{Q}_{k,N}$ is close to $\mathcal{Q}_N$ and covers X.

From this we deduce that G is a subgroup of H as follows. Given $g \in G$, take X so that $\mathcal{O}$, $g(\mathcal{O})$ lie in its interior. We know $g = \lim g_k$, $g_k \in G_k$, and for large k, $g_k(\mathcal{O}) \in X$. Therefore g_k is a word of length $\leq N$ in the face pairing transformations of $\mathcal{P}(G_k)_r$. In the limit, $g \in H$.

Now assume that H has a finite number of generators $\{h\}$. By Theorem 2.5.3, H is finitely presented. Fix a presentation. Each generator h of H is a word in the face pairing transformations of $\mathcal{P}(H)$ (centered at $\mathcal{O}$). For all sufficiently large k, say $k \geq k_0 = k_0(h)$, designate by $\psi_k(h)$ that element of G_k which is the same word in the corresponding face pairing transformations $\mathcal{P}(G_k)$. Then $\lim \psi_k(h) = h$.

The correspondence $h \mapsto \psi_k(h)$ determines a homomorphism $H \to G_k$, for large k. For if $R(h) = \mathrm{id}$ is a relation in H, we have $\lim_{k\to\infty} \psi_k(R(h)) = \mathrm{id}$, and by Lemma 4.4.1, $\psi_k(R(h)) = \mathrm{id}$ for $k \geq k_1$, where $k_1 > k_0$ is sufficiently large. Every relation in H is a consequence of the finite number in our presentation so it is only these we have to worry about. Therefore our argument shows that ψ_k determines a homomorphism as claimed.

If in addition $G = \langle g_1, \ldots, g_r \rangle$ is finitely generated, then for each index we have $g_i = \lim g_{i,k}$, with $g_{i,k} \in G_k$. By Theorem 4.1.1, the correspondence $\phi_k : g_i \to g_{i,k}$ determines a homomorphism for all large indices. Each generator g_i of G is also a word W_i in the generators $\{h\}$ of H. We know that $\lim_{k\to\infty} g_{i,k}^{-1}\psi_k(W_i) = \mathrm{id}$. Therefore for all large indices k, $g_{i,k}\psi_k(W_i) = \mathrm{id}$, that is $g_{i,k} = \psi_k(W_i)$ for all $1 \leq i \leq r$. So the homomorphism $\psi_k : H \to G_k$ is *onto* G_k; it restricts to the homomorphism $\phi_k : G \to G_k$ given by Theorem 4.1.1. □

There are many examples, in particular examples of fuchsian groups, for which polyhedral convergence does not imply algebraic convergence. Taking this into account, we note that the existence of the homomorphism $\psi_k : H \to G_k$ a few lines above did not actually require that $\{G_k\}$ have an algebraic limit. Thus:

Corollary 4.4.4. *Suppose the sequence of kleinian groups* $\{G_k\}$ *converges polyhedrally to a finitely generated kleinian group* H*. Then there is a homomorphism* ψ_k *of* H into G_k *for all large* k *such that* $\lim \psi_k(h) = h$ *for all* $h \in H$.

The following result of Brock, Bromberg, Evans and Souto is a consequence of Theorems 4.6.3 and 4.6.2(ii).

Theorem 4.4.5 [Brock et al. 2003; Brock and Souto 2006]. *Any algebraic limit of geometrically finite groups is also the geometric limit of geometrically finite groups.*

4.5 Convergence of limit sets and regions of discontinuity

Hausdorff and Carathéodory convergence

In a discussion about convergence of sequences of kleinian groups, it is natural to ask about concomitant convergence of the regions of discontinuity, or of the limit sets. The precise definitions are as follows. We begin by introducing the notion of Hausdorff convergence.

The *Hausdorff distance* between closed sets Λ and Λ_n in $\mathbb{S}^2$ is defined as follows with respect to balls $B_r(x) \subset \mathbb{S}^2$ of radius r about x in the spherical metric:

$$d_H(\Lambda, \Lambda_n) = \inf\left\{r : \Lambda \subset \bigcup_{x\in\Lambda_n} B_r(x), \text{ and } \Lambda_n \subset \bigcup_{x\in\Lambda} B_r(x)\right\}.$$

We then say that there is *Hausdorff convergence* $\lim \Lambda_n = \Lambda$ if $d_H(\Lambda, \Lambda_n) \to 0$. In words, $\lim \Lambda_n = \Lambda$ if every neighborhood of Λ contains all but a finite number of Λ_n and if U is an open set containing all but finitely many Λ_n then $\Lambda \subset U$.

The following is a standard fact about Hausdorff distance:

Lemma 4.5.1. *If* $\{\Lambda_n\}$ *is a sequence of closed sets in* $\mathbb{S}^2$*, there is a subsequence* $\{\Lambda_m\}$ *which converges in the Hausdorff topology to a closed set* $\Lambda \in \mathbb{S}^2$.

We will give two definitions of convergence of simply connected regions in $\mathbb{C}$, not the whole plane. The first assumes that the limiting region is known. In the second, the limiting region needs to be found as well. The latter is analogous to our criterion for geometric convergence of manifolds. For more details on this subject see [Duren 1983] or [Pommerenke 1992].

Situation 1. The sequence of regions $\{\Omega_n\}$ is said to converge to the region $\Omega \neq \mathbb{C}$ *in the sense of Carathéodory* if and only if every compact subset K of Ω lies in Ω_n for all large n *and* one of the following holds:

(i) Each $\zeta \in \partial\Omega$ is the limit $\zeta = \zeta_n$, $\zeta_n \in \partial\Omega_n$.
(ii) Any open set U that lies in all elements of an infinite subsequence $\{\Omega_{i_j}\}$ also lies in Ω.

Carathéodory convergence does not imply the Hausdorff convergence of the boundaries. For example, the sequence of boundaries may converge to a circle with an external ray while the regions themselves converge to the enclosed disk.

However, given a sequence of regions $\{\Omega_n\} \subset \mathbb{S}^2$, there is a subsequence such that $\{\mathbb{S}^2 \setminus \Omega_m\}$ Hausdorff converges to a closed set $\Lambda \in \mathbb{S}^2$. Then $\{\Omega_m\}$ converges in the sense of Carathéodory to $\mathbb{S}^2 \setminus \Lambda$. And conversely, if $\{\Omega_m\}$ so converges to Ω then $\{\mathbb{S}^2 \setminus \Omega_m\}$ Hausdorff converges to $\mathbb{S}^2 \setminus \Omega$.

Or more generally, $\{\partial\Omega_n\}$ Hausdorff converges to $\partial\Omega$ if and only if both $\{\Omega_n\}$ converges to Ω, and $\{\mathbb{S}^2 \setminus \overline{\Omega_n}\}$ converges to $\mathbb{S}^2 \setminus \overline{\Omega}$, both in the sense of Carathéodory.

Situation 2. Suppose $\{\Omega_k\}$ is a sequence of regions on $\mathbb{S}^2$ all of which contain a point O serving as basepoint. To avoid shrinkage to O, we will assume that a small disk about O is contained in the members of the sequence. The *kernel* of the sequence is defined to be the largest region Y containing O with the property that $\Omega_k \subset Y$ for all k with at most a finite number of exceptions. More precisely, let Y_n denote the component of $\operatorname{Int}\big(\bigcap_{k\geq n} \Omega_k\big)$ that contains O. Then $Y = \cup_n Y_n$.

A sequence $\{\Omega_n\}$ converges in the sense of Carathéodory to its kernel Y if and only if every infinite subsequence also has Y as its kernel. If Y has a hyperbolic metric, it is the limit of hyperbolic metrics on the approximating regions, uniformly on compact subsets of Y.

The kernel very much depends on the choice of basepoint O. For example, a sequence of simply connected regions may pinch in half, resulting in convergence, say, to the union of two disks. Depending on where the basepoint is chosen, the Carathéodory limit will be one or the other of the disks.

Carathéodory Convergence Theorem 4.5.2. *Suppose that $\{\Omega_n\}$ is a sequence of simply connected regions which converge in the sense of Carathéodory,* $\lim \Omega_n = \Omega$, *with respect to the basepoint $O \in \cap\Omega_n$. Assume that $\partial\Omega \subset \mathbb{S}^2$ contains at least two points. Let $f_n : \mathbb{D} \to \Omega_n$ be the Riemann map normalized by $f(0) = O$, $f'(0) > 0$. Then the sequence $\{f_n\}$ converges, uniformly on compact subsets of $\mathbb{D}$, to the normalized Riemann map $f : \mathbb{D} \to \Omega$.*

As a consequence, the sequence of hyperbolic metrics converges to the hyperbolic metric on the limiting region Ω.

Multiply connected regions can also be examined by normalizing the fuchsian covering groups with respect to O and then examining the groups with respect to geometric convergence. See Exercise 4-8.

More generally, the curvature of the metrics can be allowed to increase to 0 so the metric becomes euclidean as degeneration occurs. This results in other kinds of geometric limits.

Sequences of limit sets and regions of discontinuity

Proposition 4.5.3. *Assume that Γ is finitely generated and $\varphi_n : \Gamma \to G_n$ is a sequence of homomorphisms onto kleinian groups G_n which converges algebraically to a kleinian group G with $\Omega(G) \neq \varnothing$ and geometrically to H. Suppose that $\{\Omega(G_n)\}$ converges to $\Omega(G)$ in the sense that any given compact subset $K \subset \Omega(G)$ satisfies $K \subset \Omega(G_n)$ for all large indices. Then $\Omega(G_n)$ converges to $\Omega(G)$ in the sense of*

Carathéodory and $\Lambda(G_n)$ converges to $\Lambda(G)$ in the sense of Hausdorff. Moreover G has finite index in H.

If in addition the $\{\varphi_n\}$ are isomorphisms, then $H = G$.

What is meant here is that each component Y of $\Omega(G)$ is the Carathéodory limit of components Y_n of $\Omega(G_n)$, and conversely every sequence of components Y_n of $\Omega(G_n)$ contains a subsequence which converges to a component Y of $\Omega(G)$, in the sense of Carathéodory. Each of the components is governed by its stabilizer. There are only a finite number of conjugacy classes of component stabilizers in each G_n, G (Ahlfors Finiteness Theorem).

Proof. (See [Jørgensen and Marden 1990].) Suppose H properly contains G. Then there exists $h \in H,\ h \notin G$ such that $h = \lim g_n,\ g_n \in G_n$. Select compact sets K and K' such that $K \subset \mathrm{Int}(K') \subset K' \subset \Omega(G)$. The sequence $\{g_n(K)\}$ converges to $h(K)$. We claim that $h(K) \subset \Omega(G)$.

If not, the interior $\mathrm{Int}\, h(K') = \lim g_n(\mathrm{Int}\, K')$ contains limit points of G, in particular fixed points of loxodromic elements of G. It contains a fixed point of a loxodromic element $\varphi(\gamma) = \lim \varphi_n(\gamma)$ for some $\gamma \in \Gamma$. For all large n, $\mathrm{Int}\, g_n(K')$ contains a fixed point of $\varphi_n(\gamma)$, in contradiction to our hypothesis.

Consequently $h(K) \subset \Omega(G)$ for every compact subset K of $\Omega(G)$ and every $h \in H$. Therefore $h(\Omega(G)) \subset \Omega(G)$. The same argument can be applied to h^{-1}. We conclude that each $h \in H$ maps $\Omega(G)$ onto itself. In particular the fixed points of all loxodromic and parabolic elements of H lie in the limit set $\Lambda(G)$ showing that $\Lambda(H) = \Lambda(G)$.

In particular every fixed point of H is the limit of fixed points of G_n. Therefore every limit point of H is the limit of fixed points of G_n. We conclude that $\Omega(G_n)$ converges in the sense of Carathéodory to $\Omega(G)$.

Furthermore, if an open set $U \supset \Lambda(G)$, then also $U \supset \Lambda(G_n)$ for large n. If instead $U \supset \Lambda(G_n)$ for all large indices, then $U \supset \Lambda(G)$. Therefore the limit sets converge in the Hausdorff topology.

At this point we bring back [Greenberg 1974] (see Exercise 3-14) which implies that if $\Lambda(G)$ is not a round circle in $\mathbb{S}^2$, then G, which we know is contained in H, has finite index in H. This holds even when $\Lambda(G)$ is a circle. For G is then a fuchsian group of finite area, or a $\mathbb{Z}^2$-extension of one (via an order two elliptic), so the larger discrete group H must contain G as a subgroup of finite index ($\mathbb{H}^2/G$ is necessarily a finite-sheeted covering surface of $\mathbb{H}^2/H$).

Now we come to the assumption that each φ_k is an isomorphism. In this case we claim that $H = G$. For suppose there were an element $h = \lim \varphi_n(\gamma_n),\ \gamma_n \in \Gamma,\ h \notin G$.

Case 1: h is not elliptic. Since G has finite index in H, for some m, $h^m \neq \mathrm{id} \in G$ and $h^m = \varphi(\beta),\ \beta \in \Gamma$. Therefore $\lim \varphi_n(\gamma_n^m \beta^{-1}) = \mathrm{id}$. By Lemma 4.4.1, $\beta = \gamma_n^m$ for all large n. In a discrete group an element of infinite order has fewer than m m-th roots, by Lemma 1.5.2. Therefore for a subsequence, we can assume all γ_n are the same and $h \in G$, a contradiction.

Case 2: h is elliptic. Choose a loxodromic element $g \in G$ whose fixed points are not interchanged by h. By a direct computation using a standard form for g, we see that for some integer m, $g^m h$ is not elliptic. Case 1 again applies to show that $g^m h \in G$ and hence $h \in G$, a contradiction. □

If $\{G_n\}$ converges geometrically to H, and H is geometrically finite, that is, if the fundamental polyhedron $\mathcal{P}$ for H at any suitable basepoint has a finite number of faces, then we can deform $\mathcal{P}$ backwards. That is, suppose the face pairing, edge pairing, and vertex pairing transformations associated with $\mathcal{P}$ are moved back to G_n. Just using this finite set of elements in G_n for large n form the corresponding Dirichlet region $\mathcal{P}_n^*$. One can show that $\mathcal{P}_n^* = \mathcal{P}_n$, the fundamental polyhedron for G_n. Using this idea, as in [Jørgensen and Marden 1990], one concludes that $\{\Omega(G_n)\}$ converges to $\Omega(H)$ in the sense of Carathéodory. In view of polyhedral convergence Proposition 4.3.2, this argument leads to:

Theorem 4.5.4. *Suppose $\theta_n : \Gamma \to G_n$ is a sequence of isomorphisms of a group Γ onto kleinian groups G_n that converges algebraically to $\theta : \Gamma \to G$. Suppose G is geometrically finite with $\Omega(G) \neq \varnothing$. Then $\{G_n\}$ converges geometrically to G if and only if the regions of discontinuity converge, $\Omega(G_n) \to \Omega(G)$, in the sense of Carathéodory or equivalently, if and only if $\Lambda(G_n)$ converges to $\Lambda(G)$ in the sense of Hausdorff.*

The definitive statement of limit set convergence is due to R. Evans and is as follows; its full proof uses Theorem 5.1.2(ii), p. 242, and numerous prior results (Exercise 4-6). Note that there is no assumption about parabolics.

Theorem 4.5.5 [Evans $\geq$ 2007; Evans 2006]. *Suppose $\{\theta_n : \Gamma \to G_n\}$ is a sequence of isomorphisms from a geometrically finite group Γ to groups G_n, not necessarily geometrically finite. Assume that the sequence converges algebraically to $\theta : \Gamma \to G$ and geometrically to H. Then $\lim \Lambda(G_n) = \Lambda(H)$, in Hausdorff convergence.*

The sequence converges geometrically to G if and only if $\lim \Lambda(G_n) = \Lambda(G)$.

4.6 New parabolics

In the example of Section 4.9, a sequence of cyclic loxodromic groups converges algebraically to a cyclic parabolic group and geometrically to a rank two parabolic group. In particular the algebraic limit acquires a "new" parabolic.

More generally, if $\theta_n : \Gamma \to G_n$ is a sequence of isomorphisms converging algebraically to the isomorphism $\theta : \Gamma \to G$, then we say $g \in G$ is a *new parabolic* if for all large indices, $\theta_n \theta^{-1}(g)$ is not parabolic. We may assume that the sequence also has a geometric limit $H \supset G$.

It was conjectured by Troels Jørgensen that if $\Omega(G) \neq \varnothing$ then $H = G$ provided G does not contain new parabolics (the converse is not true). When $\Omega(G) = \varnothing$, he conjectured that always $H = G$, since there is no "room" for new elements to appear. Both of these conjectures have been confirmed, as indicated below.

Here is the description of what happens in the geometrically finite cases.

Theorem 4.6.1 [Jørgensen and Marden 1990]. *Suppose that Γ is a finitely generated abstract group without elements of finite order and $\{\theta_n : \Gamma \to G_n\}$ a sequence of isomorphisms onto kleinian groups that converges algebraically to $\theta : \Gamma \to G$. Assume that $\{G_n\}$ converges geometrically to a geometrically finite group H with $\Omega(H) \neq \varnothing$. Then*:

(*i*) *The limit sets converge* $\lim \Lambda(G_n) \to \Lambda(H)$ *in the Hausdorff topology and the sets of discontinuity converge*, $\Omega(G_n) \to \Omega(H)$, *in the sense of Carathéodory*.
(*ii*) *G is also geometrically finite*.
(*iii*) *For all large n, there is a homomorphism $\psi_n : H \to G_n$ such that* $\lim \psi_n(h) = h$ *for all $h \in H$ and for $g \in G$, $\psi_n(g) = \theta_n\theta^{-1}(g)$*.
(*iv*) *Let $\{P_j\}$, $1 \leq j \leq N$, denote the rank two parabolic subgroups of H for which $\psi_n(P_j)$ is cyclic loxodromic, one representative from each conjugacy class in H. Let $T_{j,n} \in H$ denote a generator of the kernel of $\psi_n : P_j \to \psi_n(P_j)$. Then* $\mathrm{Ker}(\psi_n)$ *is the normal closure in H of the subgroup generated by $\{T_{j,n}\}$, $1 \leq j \leq N$*.
(*v*) *Assume each P_j contains an element of G. Then there exists $T_j \in P_j$, $T_j \notin G$ such that*

$$H = \langle G, T_1, T_2, \ldots, T_N \rangle.$$

(*vi*) *$H = G$ if and only if the class $\{P_j\}$ is empty*.

Outline of proof. As a finitely generated subgroup of the geometrically finite group H with $\Omega(H) \neq \varnothing$, G is also geometrically finite (Lemma 3.6.3). Item (i) follows from the remarks preceding Theorem 4.5.4. The first part of (iii) comes from Theorem 4.4.3. The second part of (iii) is a consequence of Lemma 4.4.1, namely $\theta_n\theta^{-1}(g) = \psi_n(g)$ for all large n, first for a set of generators of G and then for all G. Item (iv) is proved by working backward from a fundamental polyhedron for H to fundamental polyhedra for its approximates G_n. We will omit the detailed proof of this. The proof of (v) begins with the fact that the common fixed point ζ_j of P_j is also a parabolic fixed point of G. Once again this is established by working backwards from a fundamental polyhedron for H; $\psi_n(P_j)$ represents a simple, short geodesic in $\mathcal{M}(G_n)$ which is associated with a word of uniformly bounded length in the face pairing transformations for G_n. So let $S_j \in G$ be a generator of the parabolic subgroup that fixes ζ_j. Then $\psi_n(S_j)$ is a generator of $\psi_n(P_j)$. Consequently $P_j = \langle S_j, T_{j,n}\rangle$. Take T_j to be any one of the $T_{j,n}$.

Item (vi) requires the elementary fact that the geometric limit of an algebraically convergent sequence of cyclic parabolic groups (which is again a cyclic parabolic group) is the same as the algebraic limit. The only way that H can differ from G is that there exist rank two groups $P_j \in H$ that are geometric limits of necessarily cyclic loxodromic subgroups of $\{G_k\}$ (while their algebraic limit is a cyclic parabolic group). □

Theorem 4.6.1 has been greatly generalized through the efforts of several authors, particularly Anderson, Brock, Bromberg, Canary, Evans, Ohshika, and Souto. Here is a statement of the final result incorporating the Tameness Theorem that confirms Jørgensen's Conjecture.

Theorem 4.6.2. *Suppose* $\{\theta_n : \Gamma \to G_n\}$ *is a sequence of isomorphisms converging algebraically to* $\theta : \Gamma \to G$. *The sequence also converges geometrically to* G *under one of the following situations*:

(i) [Anderson and Canary 1996b; Evans 2004a] *If* $\Omega(G) \neq \varnothing$ *and* G *has "no new parabolics", that is,* $g \in G$ *is parabolic if and only if* $\theta_n\theta^{-1}(g)$ *is parabolic for all large indices* n.

(ii) [Canary 1996, Theorem 9.2; Agol 2004; Calegari and Gabai 2004] *If* $\Omega(G) = \varnothing$.

Theorem 4.6.2 does not require that the approximating groups be geometrically finite (just finitely generated and torsion free). Of course the converse to (i) does not hold, convergence to G can be geometric even in the presence of new parabolics. Condition (ii) was initially established under additional assumptions, in particular when G is known to be tame. By incorporating the Tameness Theorem, we can make the general statement given here. In this case, whether or not there are new parabolics makes no difference.

A sequence is often said to be *strongly convergent* if it converges both algebraically and geometrically to the limiting group.

Here is another useful fact (especially in the context of the Density Theorem on p. 260):

Theorem 4.6.3 [Brock et al. 2003]. *If* H *is the algebraic limit of geometrically finite groups, then* H *is also the algebraic limit of geometrically finite groups* $\{\theta'_n : \Gamma' \to G'_n\}$ *with the property that* $\theta'(g) = \lim \theta'_n(g)$ *is parabolic if and only if* $\theta'_n(g)$ *is parabolic for all indices*.

By Theorem 4.6.2 H is also the geometric limit of $\{G'_n\}$. Of course in general, the groups Γ, Γ' will not lie in the same quasiconformal deformation space. What is remarkable about the theorem is that there is no requirement that H be geometrically finite.

4.7 Acylindrical manifolds

A compact 3-manifold with boundary M^3 is called *acylindrical* (or anannular) if M^3 contains no essential cylinders *and* is boundary incompressible. We recall from Section 3.7 that an *essential cylinder* C in M^3 is a cylinder C such that $C \cap \partial M^3 = \partial C$ and C is not homotopic into ∂M^3.

There are two ways to apply this definition in a geometrically finite $\mathcal{M}(G)$. The usual definition, given in Section 3.7, is to call $\mathcal{M}(G)$ acylindrical if it is boundary incompressible and every essential cylinder is homotopic into $\partial\mathcal{M}(G)$ or into a pairing cylinder. This means that every component of $\Omega(G)$ is simply connected, that a loxodromic element can preserve at most one component of $\Omega(G)$, while a rank one parabolic arises only from simple loops on $\partial\mathcal{M}(G)$ retractable to its associated pair of punctures.

The second and less commonly used sense of the term is to define "acylindrical" with respect to the compact $\mathcal{M}_0(G) = \mathcal{M}(G)^{\text{thick}}$, which results from removing the interiors of solid pairing tubes and cusp cylinders. Even if $\mathcal{M}(G)$ is boundary incompressible, $\partial\mathcal{M}_0(G)$ may not be so, as removing the totality of the solid pairing tubes may bring in a new topology to the boundary. Moreover essential cylinders in $\mathcal{M}_0(G)$ do not necessarily correspond to essential cylinders in $\mathcal{M}(G)$: It is possible that there is an essential cylinder C in $\mathcal{M}_0(G)$ with the property that one component of ∂C lies on a cusp torus (Exercise 4-21).

In any case, in $\mathcal{M}(G)$ a simple (nontrivial) loop on a cusp cylinder associated with a rank one cusp or on a cusp torus cannot be freely homotopic to a simple loop on a different cusp cylinder or cusp torus, for the corresponding parabolic subgroups belong to distinct conjugacy classes.

A cyclic subgroup corresponding to an essential cylinder C is either loxodromic or parabolic. In the parabolic case, since C cannot serve as a pairing cylinder, at most one component of ∂C can be retractable in the boundary to a puncture (Exercise 4-21). Consider a component γ of ∂C which is not retractable to a puncture. Examine the component $\Omega \subset \Omega(G)$ that contains a lift γ^* of γ. The simple arc γ^* has both its endpoints at a parabolic fixed point. When the fixed point is added, γ^* becomes a Jordan curve, necessarily separating $\Lambda(G)$ into two parts. In particular $\partial\Omega$ is not a Jordan curve. (This is an example of an "accidental parabolic" transformation.)

This is a good place to interject that one way to exclude accidental parabolics is to require that for each component $\Omega \subset \Omega(G)$, the subgroup $\operatorname{Stab}(\Omega)$ is quasifuchsian (Exercise 3-10).

Acylindrical manifolds have compact algebraic deformation spaces. More precisely:

Thurston Compactness Theorem [Thurston 1986b]. *Let G be a geometrically finite group such that $\mathcal{M}(G)$ is acylindrical with nonempty boundary. Then every sequence of parabolic preserving isomorphisms to kleinian groups $\theta_n : G \to G_n$ has an algebraically convergent subsequence.*

Suppose there were, in a geometrically finite manifold $\mathcal{M}(G)$, an essential cylinder C corresponding to the conjugacy class of a cyclic loxodromic subgroup. Suppose for example C divides $\mathcal{M}(G)$ into two components M, M'. Focus on M and fix a lift $M^* \subset \mathbb{H}^3 \cup \Omega(G)$. Normalize things so that a given point $O \in \mathbb{H}^3$ lies in a ball about O in M^*. Set $G_1 = \{g \in G : g(M^*) = M^*\}$. ($\pi_1(\mathcal{M}(G))$ is the free product of the fundamental groups of M, M' amalgamated over the common cyclic subgroup determined by C.)

Then we should expect that there is a sequence of deformations of $\mathcal{M}(G_1) \cong M$ so that each cyclic loxodromic subgroup determined by C converges to a cyclic parabolic subgroup and C becomes a cusp cylinder in the limit. Here we keep the same normalization with respect to O.

If the lift (that is, a component of the preimage) M'^* of M' is adjacent to M^* and $G_2 = \text{Stab}(M'^*)$, then except for the cyclic subgroup of G_2 that corresponds to the common boundary with M^*, the group G_2 will simply disappear in the limit—the Möbius transformations do not converge. This is why the acylindrical condition is necessary in the Compactness Theorem. Such phenomena appear in particular for fuchsian groups, see Exercise 4-8. One can start with a fuchsian group Γ and the lift of a simple geodesic from $\mathbb{H}^2/\Gamma$, and "pinch" the geodesic so that in the limit it corresponds to a parabolic transformation. Thurston's theorem says that such degenerations are impossible if there are no essential cylinders to begin with.

Supplementing the Thurston Compactness Theorem we have:

Theorem 4.7.1 [Johannson 1979; Matsuzaki and Taniguchi 1998, Theorem 3.29]. *Suppose G is geometrically finite such that $\mathcal{M}(G)$ is acylindrical with nonempty boundary. Let $\theta : G \to G'$ be an isomorphism to a geometrically finite G' such that $\theta(g)$ is parabolic if and only if $g \in G$ is parabolic. Then there exists a quasiconformal mapping $F : \mathbb{S}^2 \to \mathbb{S}^2$ that satisfies $F \circ g \circ F^{-1}(z) = \theta(g)(z)$ for all $g \in G,\ z \in \mathbb{S}^2$. It can be chosen to project and extend to be a (quasiisometric) homeomorphism $F_* : \mathcal{M}(G) \to \mathcal{M}(G')$.*

If an initial mapping F turns out to be orientation reversing, it can be replaced by JF and G' by $JG'J$ where J is reflection in a plane in $\mathbb{H}^3$. Of special interest is the fact that θ dictates a bijection between components of $\Omega(G')$ and $\Omega(G)$. See Exercise 4-9.

4.8 Dehn surgery

Dehn surgery is an operation performed one or more incompressible torus boundary components of a manifold M^3. Choose a torus boundary component $\mathbb{T}$ and a pair of simple loops α, β, crossing each other once, so as to generate its homology and homotopy. Once α, β are chosen, (the homology class of) every simple loop γ on the torus can be expressed the form $\gamma = m\alpha + n\beta$ where m, n are relatively prime integers. The ratio $0 \le n/m \le \infty$ is called its *slope* (in terms of the choice of α, β).

Choose such a simple curve $\gamma = m\alpha + n\beta$, not homologous to zero, on $\mathbb{T}$. Glue to M^3 along $\mathbb{T}$ a solid torus in such a way that γ becomes a *meridian*, that is, γ bounds a disk in the new solid torus. This is the process of (m, n)-Dehn surgery; the designation of a simple loop γ on $\mathbb{T}$ as a meridian tells us how to add a solid torus to the boundary component to result in a larger manifold—with one less boundary component. Dehn surgery can be applied to any or all of the torus boundary components.

Another way to describe γ is as the image of α under an automorphism ϕ of $\mathbb{T}$. Gluings by two automorphisms ϕ, ϕ' determine homeomorphic manifolds if and only if $\phi_1 \circ \phi^{-1}$ extends to a homeomorphism between the solid tori.

A more typical implementation is as follows. Choose a simple loop, or a number of mutually disjoint simple loops in the interior of a 3-manifold M^3, for example a link in $\mathbb{S}^3$. Enclose the loops by mutually disjoint tubular neighborhoods. Unlike the case for a cusp torus, in this situation each torus boundary has a uniquely determined (up to free homotopy) meridian α, that bounds a disk in the solid torus tubular neighborhood. Choose a simple loop β that crosses α exactly once and with α generates the homology of $\mathbb{T}$. Now choose a simple loop $\gamma = m\alpha + n\beta$. Remove the tubular neighborhood bounded by $\mathbb{T}$ and replace it by gluing in a new solid torus so that γ becomes its meridian. This process can be applied to each of the tubular neighborhoods.

For a hyperbolic manifold $\mathcal{M}(G)$ with a rank two cusp, the process can be applied to a cusp torus $\mathbb{T}$ and a pair of generators α, β of its homology. If the cusp torus arises from $\langle z \mapsto z+1,\ z \mapsto z+\tau\rangle$, we may choose α to correspond to the first generator and β the second. Then choose the simple loop γ corresponding to $m+n\tau$. Remove the solid cusp torus bounded by $\mathbb{T}$ and replace it by a solid torus in terms of which γ becomes a meridian. This gives a new manifold M^3 in which the cusp torus becomes a tubular neighborhood of a nontrivial simple loop — but the initial hyperbolic structure is lost.

The latter operation are commonly called *Dehn filling*.

4.9 The prototypical example

This is an explicit example both of Dehn surgery in the simplest case and of differing algebraic and geometric limits. We will start with a solid cusp torus — a rank two parabolic group — and do $(1, n)$ Dehn surgery on it. There results a cyclic loxodromic group. We will then watch what happens as $n \to \infty$. Figure 1.5 (p. 12) and Figure 4.1 (p. 209) show several generations of isometric circles of a cyclic loxodromic group.

Start with the parabolic group $\Gamma = \langle T_1(z) = z + \omega_1,\ T_2(z) = z + \omega_2\rangle$. Set $\tau = \omega_2/\omega_1,\ \operatorname{Im}\tau > 0$. The quotient $\mathbb{C}/\Gamma = \mathcal{T}$ is a torus. The generating pair (ω_1, ω_2) corresponds to a pair of simple loops α, β on $\mathcal{T}$, crossing each other once.

Change the basis by the rule

$$\omega_{1,n} = \omega_1 + n\omega_2, \quad \omega_{2,n} = \omega_2,; \quad \tau_n = \frac{\omega_{2,n}}{\omega_{1,n}} = \frac{\tau}{1+n\tau},$$

so that $T_{1,n}(z) = z + \omega_{1,n},\ T_{2,n}(z) = z + \omega_{2,n}$ also generate Γ. The pair $(\omega_{1,n}, \omega_{2,n})$ represents the simple loops $\alpha + n\beta,\ \beta$ on $\mathcal{T}$.

Map $\mathbb{C}$ onto $\mathbb{C}\setminus\{0\}$ by $w_n(z) = e^{-2\pi i z/\omega_{1,n}}$.

Let U_n denote the loxodromic transformation

$$U_n(w) = e^{-2\pi i \tau_n}\, w = a_n w.$$

We have $(w_n \circ T_1)(z) = ({U_n}^{-n} \circ w_n)(z)$ and $(w_n \circ T_2)(z) = (U_n \circ w_n)(z)$, while $(w_n \circ T_{1,n})(z) = w_n(z)$ and $(w_n \circ T_{2,n})(z) = (U_n \circ w_n)(z)$.

The map w_n determines a conformal mapping

$$\mathcal{T} \to \mathcal{T}_n = (\mathbb{C}\setminus\{0\})/\langle U_n\rangle$$

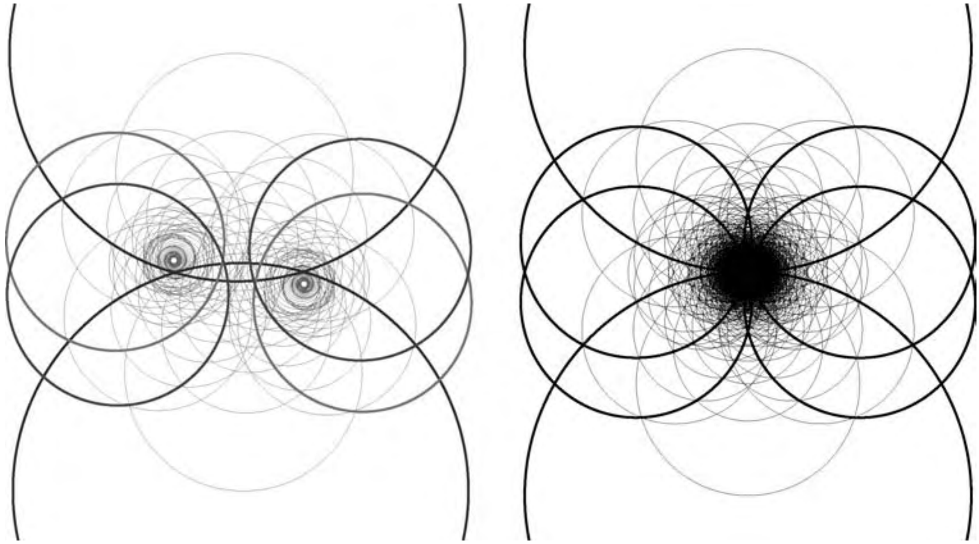


Fig. 4.1. A cyclic group generated by a loxodromic of approximate trace $1.919354 + 0.029772i$ near its rank-2 parabolic geometric limit in the right frame. One can see how the 6-sided Ford polygon outside the outer circles is becoming a fundamental domain on $\mathbb{S}^2$ for the geometric limit. See [Jørgensen 1973] for a description of the combinatorics of the approximates.

in which the image of $\alpha + n\beta$ is a meridian in the solid torus $(\mathbb{H}^3 \cup (\mathbb{C} \setminus \{0\}))/\langle U_n \rangle$. The image of straight lines with tangent vector $\omega_{1,n}$ are taken by w_n to concentric circles about $w = 0$ which in turn project to parallel meridians in $\mathcal{T}_n$. We have done Dehn surgery on the original cusp torus $\mathcal{M}(\Gamma)$ by removing $\mathrm{Int}(\mathcal{M}(\Gamma))$ and replacing it by a solid torus so the chosen simple loop $\alpha + n\beta$ becomes a meridian.

As $n \to \infty$, $\lim \tau_n = 0$, $\lim a_n = 1$, and $\lim U_n = \mathrm{id}$. Renormalize U_n to have the fixed points $\omega_2/(1 - a_n)$, ∞, thus

$$A_n(w) = w + \frac{\omega_2}{1 - a_n}, \qquad V_n(w) = A_n U_n A_n^{-1}(w) = a_n w + \omega_2.$$

Therefore $\lim V_n(w) = w + \omega_2$ and

$$V_n{}^k(w) \ = \ A_n U_n^k A_n^{-1}(w) \ = \ a_n^k\, w + \frac{a_n^k - 1}{a_n - 1}\, \omega_2.$$

Define

$$f_n(z) = \ \frac{\omega_2}{a_n - 1}\, (w_n(z) - 1).$$

Thus $f_n \circ T_1(z) = V_n{}^{-n} \circ f_n(z)$ and $f_n \circ T_2(z) = V_n \circ f_n(z)$ while $f_n \circ T_{1,n}(z) = f_n(z)$ and $f_n \circ T_{2,n}(z) = V_n \circ f_n(z)$. In short, f_n induces a conformal mapping to the renormalized solid torus

$$\mathcal{T} \to \left(\mathbb{C} \setminus \left\{\frac{\omega_2}{1 - a_n}\right\}\right)/\langle V_n \rangle.$$

in such a way that the image of $\alpha + n\beta$ remains a meridian.

Uniformly on compact subsets of $\mathbb{C}$ we have the following convergences:

(i) $\lim f_n(z) = z$.
(ii) $\lim V_n(w) = w + \omega_2$.
(iii) $\lim V_n{}^{-n}(w) = w + \omega_1$.

To prove (i), we use the estimate $e^x - 1 \sim x$ when x is small, and (iii) follows.

It is more complicated to show:

Claim. *The sequence of cyclic loxodromic groups $\{\langle V_n \rangle\}$ converges algebraically to the cyclic parabolic group $\langle T_2 \rangle$ and geometrically to the rank two parabolic group $\Gamma = \langle T_1, T_2 \rangle$.*

Proof. Suppose for a sequence $m \to \infty$ and $k = k(m) \to \infty$ that $\{V_m{}^k\}$ converges to a Möbius transformation. We must show the limit lies in Γ. For the limit to exist the ratio $(a_m{}^k - 1)/(a_m - 1)$ must remain bounded. Therefore $\lim a_m{}^k = 1$.

Write $k = pm + q$ where p, q are integral functions of m, as is k, and $0 \le q < m$. Since $a_m{}^k = \exp(-2\pi i k \tau_m)$ and $\operatorname{Im} \tau_m = \operatorname{Im} \tau / |1 + m\tau|^2$, we must have $k(m) = o(m^2)$. Therefore $p(m) = o(m)$, as $m \to \infty$.

Take the subsequence $\{m\}$ so that $\lim q(m)/m = c$ exists, $0 \le c \le 1$. We claim that either $c = 0$ or $c = 1$. For first of all $e^{-2\pi i k \tau_m} = e^{-2\pi i (k\tau_m - p)} = e^{-2\pi i (pm\tau_m - p + q\tau_m)}$. Also,

$$\lim\, p(m\tau_m - 1) = -\lim \frac{p}{1 + m\tau} = 0, \qquad \lim\, q\tau_m = c.$$

So if c were not an integer, the ratio

$$\frac{a_m^k - 1}{a_m - 1} = \frac{e^{-2\pi i k \tau_m} - 1}{e^{-2\pi i \tau_m} - 1} \tag{4.1}$$

would become infinite.

We have to examine (4.1) in more detail. Write

$$e^{-2\pi i k \tau_m} = e^{-2\pi i (k\tau_m - p - c)}$$

so that the exponent approaches zero as $m \to \infty$. By Taylor's formula for e^x, the limit of the ratio (4.1) is the limit of

$$\frac{k\tau_m - p - c}{\tau_m} = \frac{p(m\tau_m - 1) + q\tau_m - c}{\tau_m} = \frac{-p + (q - cm)\tau - c}{\tau}.$$

Since $\operatorname{Im} \tau > 0$, if this is to have a finite limit then $\lim_{m\to\infty}(q - cm)$ must exist, necessarily as an integer. Then $\lim_{m\to\infty} p$ must exist as well, also as an integer. We conclude that

$$\lim_{m\to\infty} V_m^k(w) = w - \omega_1(c + \lim p) + \omega_2 \lim(q - cm). \qquad \square$$

Summing up, the solid tori $\mathcal{M}(\langle V_n \rangle)$ converge algebraically to $\mathcal{M}(\langle T_2 \rangle)$, which represents a solid cusp tube associated with a cyclic parabolic group. The boundary torus has become "pinched" — has become a doubly infinite cylinder: the hole in the bagel has coalesced to a single point since the length of the geodesic has gone to zero. In contrast, the manifolds $\mathcal{M}(\langle V_n \rangle)$ converge geometrically to the solid cusp torus $\mathcal{M}(\Gamma)$. The process of degeneration introduces so much twisting along the boundary torus, that in the limit the solid torus fractures, with the fracture line lying at its core. All this happens while the conformal type of the torus itself does not change, what changes is the presentation of its fundamental group.

For a generic choice of center $O \in \mathbb{H}^3$, the fundamental polyhedron $\mathcal{P}$ for Γ is a 6-sided chimney rising from $\mathbb{C}$. The approximates $\mathcal{P}_n$ acquire more and more faces as $n \to \infty$, but all of the faces, save six, collapse to the fixed point in the limit. The polyhedra truncated by intersection with a ball of radius r about their center $\mathcal{P}_{r,n}$ converge uniformly to $\mathcal{P}_r$.

Remark 4.9.1. We can now give an example of a sequence of cyclic loxodromic groups $\{\langle S_m \rangle\}$ that converge geometrically to a rank two parabolic group, yet which do not converge algebraically. In fact no subsequence of the generators $\{S_m\}$ has a limit.

For an example, take from above $V_m(z) = a_m z + \omega_2$ with $\lim a_m = 1$. Pick any sequence of integers $n = n(m)$ which go to infinity with m. Set $d_m = \sqrt[n]{a_m}$ where the root is chosen in any way so long as *no* subsequence approaches 1. Set

$$S_m(z) = d_m z + \frac{d_m - 1}{d_m^n - 1}\omega_2.$$

Then $S_m^n(z) = V_m(z)$ and $S_m^{nk}(z) = V_m^k(z)$ but no subsequence of $\{S_m\}$ converges since the constant term $\to \infty$.

At the other extreme a sequence of cyclic loxodromic groups $\langle z \mapsto c_n z \rangle$ with $c_n > 0$, $c_n \to 1$ can be conjugated to converge algebraically and geometrically to a cyclic parabolic group. Only loxodromics whose traces converge "tangentially" to ± 2 can have differing algebraic and geometric limits.

The space of cyclic loxodromic groups is completely described in terms of the combinatorics of the faces of the Ford polyhedron in [Jørgensen 1973] and visualized in Wada's program [$\geq$ 2007a], which allows exploration of the space of normalized cyclic loxodromic groups as a function of the trace.

4.10 Manifolds of finite volume

Suppose $\{\mathcal{M}(G_n)\}$ is an infinite sequence of mutually nonisometric manifolds whose volumes $\{V_n\}$ do not exceed a number $V^* < \infty$. We can normalize the groups so that an ε-ball centered at a point $\mathcal{O} \in \mathbb{H}^3$, projects injectively into all of the manifolds. After passing to a subsequence, we may assume that the sequence $\{G_n\}$ converges geometrically to a group H.

Theorem 4.10.1. *Assume that the sequence $\{G_n\}$ with $\mathrm{Vol}(\mathcal{M}(G_n)) = V_n < V^* < \infty$ converges geometrically to H. Then $\mathrm{Vol}(\mathcal{M}(H)) = \lim V_n$. Consequently the set of volumes of finite volume manifolds is a closed subset of $\mathbb{R}$.*

Moreover, the number of solid cusp tori in $\mathcal{M}(H)$ is strictly greater than the number in its approximates, for large indices.

Proof. Take a maximal set of points in the ε-thick part $\mathcal{M}(G_n)^{\text{thick}}$ such that the distance between any two of them is not less than $\varepsilon/2$. Then the ε-balls about these points cover $\mathcal{M}(G_n)^{\text{thick}}$. The fact that $\text{Vol}(\mathcal{M}(G_n) < V^*$ implies that the number of the covering balls is uniformly bounded in n. In particular there exists $d < \infty$ such that for all indices, the diameter of $\mathcal{M}(G_n)^{\text{thick}}$ does not exceed d. As $n \to \infty$, $\text{Vol}(\mathcal{M}(G_n)^{\text{thick}})$ converges to $\text{Vol}(\mathcal{M}(H)^{\text{thick}})$. Exercise 2-9, especially Equation (2.6), shows that $\lim_{\varepsilon\to 0} \text{Vol}(\mathcal{M}(G_n)^{\text{thin}}) = 0$, uniformly in n—the thin parts become successively thinner. This shows that $\text{Vol}(\mathcal{M}(H)) = \lim \text{Vol}(\mathcal{M}(G_n))$; in particular $\mathcal{M}(H)$ has finite volume.

Now that we know $\mathcal{M}(H)$ has finite volume, the corresponding polyhedra for G_n must have a uniformly bounded number of faces and hence generators—see Lemma 3.6.4. The methods of the proof of Theorem 4.6.1 apply to describe the relation of the nearby polyhedra $\mathcal{P}_{\mathcal{O}}(G_n)$ to the polyhedron $\mathcal{P}_{\mathcal{O}}(H)$ for H.

For all large n there is a homomorphism $\psi_n : H \to G_n$. The ψ_n-image of each rank two parabolic subgroup of H is either a rank two subgroup of G_n or it represents the lift of a short geodesic in $\mathcal{M}(G_n)$.

If $\mathcal{M}(H)$ had the same number of solid cusp tori as $\mathcal{M}(G_n)$ for all large indices, then ψ_n would be an isomorphism and $\mathcal{M}(G_n)$ would be isometric to $\mathcal{M}(H)$, by Mostow's Rigidity Theorem, for the same large indices. Our assumption rules out this possibility. There are always strictly more solid cusp tori in the geometric limit than in the approximants. A sequence of cyclic loxodromic subgroups has become a rank two parabolic group in the geometric limit. see Theorem 4.6.1. □

4.11 The Dehn surgery theorems for finite volume manifolds

Suppose $\mathcal{M}(G)$ has finite volume and has $k \geq 1$ rank two cusps. Denote by M the compact manifold bounded by k tori $\{T_i\}$ resulting from removing the interior of k solid cusp tori. We will discuss the result of doing Dehn surgery on these. (We can allow additional rank two cusps that will then be unaffected.)

Choose a standard homology basis (γ_i, δ_i) on each T_i, $1 \leq i \leq k$.

Let $Q \subset \mathbb{S}^2$ the set of coprime vectors $\{d_i = (p_i, q_i)\}$. Given $d = (d_1, \dots, d_k) \in Q^k$, denote by M_d the manifold resulting from (p_i, q_i)-Dehn surgery on T_i, $1 \leq i \leq k$. Here the respective meridians are $\{p_i\gamma_i + q_i\delta_i\}$. The resulting manifolds M_d are *closed*.

Before stating the theorem, we will present the argument in [Thurston 1979, §5.6] that the dimension of the local deformation space of $\mathcal{M}(G)$ under the Dehn surgeries is $\geq k$. See [Culler and Shalen 1983, Prop. 3.2.1] for an alternate treatment.

For each index i, $1 \leq i \leq k$, choose in any way a simple, noncontractible loop $\{\alpha_i\}$ in the interior of $M \subset \mathcal{M}(G)$ that corresponds to a loxodromic transformation in G. Fix a basepoint $O_i \in T_i$ that is also the basepoint for α_i. Remove a thin tube about each α_i and attach it to T_i so as to form a surface S_i of genus two. Do this for all indices, assuming the α_i are mutually disjoint, ending up with a manifold $M' \subset M$ bounded by k surfaces of genus two. On each S_i chose a simple loop β_i that bounds a compressing disk in $M \setminus M'$. We can regard α_i to lie on S_i and also that the four simple loops $\gamma_i, \delta_i, \alpha_i, \beta_i$ have the common basepoint O_i. The four loops generate $\pi_1(S_i; O_i)$, and satisfy the relation $[\delta_i, \gamma_i][\beta_i, \alpha_i] = \text{id}$.

Note that $\pi_1(M) \cong G$ is obtained from $\pi_1(M')$ by adding the relations $\{\beta_i = \text{id}\}$; if we cut along the compressing disks bounded by the β_i, we obtain a manifold homeomorphic to M.

The elements γ_i, δ_i, when lifted from a fixed O_i^* over O_i determine generators γ_i^*, δ_i^* of a rank two parabolic subgroup of G. The element α_i when lifted from O_i^* determines a loxodromic element α_i^*. Under a small deformation, their traces change slightly, and also the location of the fixed points. Therefore when a homomorphism ψ is close to id, $\psi(\alpha_i^*)$ remains loxodromic with fixed points distinct from those of $\psi(\gamma_i^*)$, $\psi(\delta_i^*)$.

Lemma 4.11.1. *Suppose ψ is a homomorphism $\pi_1(M') \to \mathrm{PSL}(2,\mathbb{C})$ such that (a) $\psi(\langle\gamma_i, \delta_i\rangle) \neq \text{id}$, (b) $\psi(\alpha_i)$ is loxodromic, and (c) $\psi(\langle\alpha_i, \gamma_i, \delta_i\rangle)$ is nonelementary. Then ψ extends to a homomorphism of $\pi_1(M)$ if and only if for each index i the following two equations are satisfied*:

$$\operatorname{tr}\psi([\alpha_i, \beta_i]) = 2, \quad \operatorname{tr}\psi(\beta_i) = 2. \tag{4.2}$$

The necessity of the condition is obvious since a homomorphism of $\pi_1(M)$ must send both the commutator and the element β_i to the identity. The sufficiency is Exercise 4-5.

Now the compact 3-manifold M' can be triangulated in such a way that there is only one 0-simplex and the 1-simplices are generators of $\pi_1(M)$. The 2-simplices then generate the relations among the chosen generators. Since the manifold has a nonempty boundary, the 2-skeleton of the triangulation is a deformation retract of M'. Its Euler characteristic is then

$$\chi(M') = +1 - h + r,$$

where h is the number of generators and r the number of relations.

Moreover $\chi(M') = \chi(M) - k$ because M is obtained from M' by adding k relations. That is,

$$\chi(M) = 1 - h + r + k.$$

The ψ-image of the h generators of G arising from our construction in M must satisfy the algebraic equations corresponding to each relation. Each Möbius transformation in turn depends on 3 complex parameters. In addition Equations (4.2) must be accounted for; that gives two more equations for each torus boundary. Thus the $3h$ parameters for the ψ-image of the generators are subject to constraints and the result is that ψ has the degree of freedom given by

$$3h - 3r - 2k = -3\chi(M) + k + 3.$$

But if we rule out conjugations of the group G, we are left with the complex dimension $-3\chi(M)+k$.

A closed 3-manifold has Euler characteristic zero. Therefore if $\widehat{M}$ denotes the double of M across its boundary,

$$0=\chi(\widehat{M})=2\chi(M)-\chi(\partial M).$$

Since all the components of ∂M are tori, $\chi(M)=0$ (for a general geometrically finite kleinian manifold we would have instead $\chi(\partial M)\le 0$ and then $\chi(M)\le 0$). Thus ψ has k degrees of freedom; each rank two cusp contributes one degree.

For a rigorous study of the deformation variety, see [Kapovich 2001, Theorem 8.44].

The following result shows that there are lots of hyperbolic manifolds, independent of the criteria of the Hyperbolization Theorem (p. 324). The paper [Petronio and Porti 2000] is the current standard for a complete, rigorous proof of the first part of the following theorem. It is quite different from the one suggested in [Thurston 1979], and reflects the computational approach of SnapPea (page 234). For another approach, see [Hodgson and Kerckhoff 1998, §4].

Dehn Surgery Theorem [Thurston 1979, §5.5–8; Petronio and Porti 2000].

(i) There exists a neighborhood U of $\infty=(\infty,\dots,\infty)\in\mathbb{S}^2\times\cdots\times\mathbb{S}^2$ such that for all $d\in U\cap Q^k$, the surgered manifold M_d has a complete hyperbolic metric.

(ii) More precisely, if a finite number of coprimes $\{(p_i,q_i)\}$ are excluded for each $\{T_i\}$, $1\le i\le k$, then all remaining Dehn surgeries on $\mathcal{M}(G)$ result in complete hyperbolic manifolds.

(iii) Suppose $\lim d_n=\infty$ in U. The hyperbolic manifolds $\mathcal{M}(G_n)\equiv M_{d_n}$ converge geometrically back to $\mathcal{M}(G)$. The corresponding homomorphisms $\psi_n: G\to G_n$ converge to the identity.

In particular there are arbitrarily small deformations $\{H\}$ of G which send any or all of the rank two parabolic subgroups to cyclic loxodromic groups. The result of removing from each such $\mathcal{M}(H)$ tubular neighborhoods of its new short geodesics is homeomorphic to $\mathcal{M}(G)$.

When the number of initial cusp tori is at least two, it is not true in general that, with a finite number of possible exceptions, all surgeries on the cusps of an $\mathcal{M}(G)$ result in hyperbolic manifolds. Consider as $\mathcal{M}(G)$ the Borromean rings complement in $\mathbb{S}^3$. The $(1,0)$ surgery on one of the links results in a manifold homeomorphic to $\mathbb{S}^3$ minus two unlinked circles. This is not hyperbolic, nor is the result of any further surgery — there are noncontractible embedded spheres in the complement. [Thurston 1979, p. 5.38].

A similar process allows the construction of orbifolds where the rank-two parabolic groups are instead sent to cyclic elliptic groups with designated rotation angles.

Well ordering of volumes of hyperbolic manifolds

By the universal ball property, there is a uniform positive lower bound for all volumes. By the uniform horoball property, there is a uniform upper bound on the number of solid cusp tori in manifolds of volume $\leq V$. Here the convergence theorems 4.1.1, 4.1.2 play a central role.

Theorem 4.11.2 [Thurston 1979, Chapter 5-6, Gromov 1981b].

(i) *The set of hyperbolic* 3*-manifolds with a given volume* V *is finite.*
(ii) *If* $\mathcal{M}(G)$ *of finite volume is homeomorphic to* $\mathcal{M}(H) \setminus \alpha$, *where* $\alpha \subset \mathcal{M}(H)$ *is a simple geodesic*, *then* $\mathrm{Vol}(\mathcal{M}(G)) > \mathrm{Vol}(\mathcal{M}(H))$.
(iii) *If* $\{\mathcal{M}(G_k)\}$ *is a sequence of manifolds whose volumes are nonincreasing*,

$$\cdots \geq \mathrm{Vol}(\mathcal{M}(G_k)) \geq \mathrm{Vol}(\mathcal{M}(G_{k+1})) \geq \cdots,$$

then $\mathrm{Vol}\mathcal{M}(G_k) = \mathrm{Vol}\mathcal{M}(G_m)$ *for some* m *and all* $k > m$.
(iv) *For each constant* V *let* $\mathfrak{M}_V$ *denote the set of hyperbolic* 3*-manifolds with volume* $\leq V$. *There is a finite subset* $\mathfrak{M}_{\mathrm{moms}} \subset \mathfrak{M}_V$ *such that any* $\mathcal{M}(H) \in \mathfrak{M}_V \setminus \mathfrak{M}_{\mathrm{moms}}$ *contains a link* L *such that* $\mathcal{M}(H) \setminus L$ *is homeomorphic to some* $\mathcal{M}(G) \in \mathfrak{M}_{\mathrm{moms}}$ *and is obtained by Dehn surgery on* $\mathcal{M}(G)$; *moreover* $\mathrm{Vol}(\mathcal{M}(H)) < \mathrm{Vol}(\mathcal{M}(G))$.

In fact according to [Thurston 1979, Theorem 5.11.2], there is a link $L_V \subset \mathbb{S}^3$ such that all manifolds in $\mathfrak{M}_V$ can be obtained by Dehn surgery along L_V (the limiting case of simply deleting components of L_V is allowed).

Heuristic discussion. The first item stems from the following argument. If there is an infinite sequence of nonisometric manifolds of volume exactly V there is a geometric limit of a subsequence. It must have at least one additional cusp which raises the volume by (ii).

We refer to [1979, Chapter 6] for the proof of (ii), that is, $\mathrm{Vol}(\mathcal{M}(H)) < \mathrm{Vol}(\mathcal{M}(G))$ when $\mathcal{M}(H) \setminus \cup\gamma_i$ is homeomorphic to $M(G)$ for a union of mutually disjoint nontrivial simple loops γ_i. The proof is based on the analysis of the volumes of hyperbolic manifolds which are the images under degree $d \geq 1$ maps of a given finite volume manifold.

The most problematical issue in the background is to prove that the number of homeomorphism types for ε-thick parts M of hyperbolic manifolds of volume at most V is finite. Here is Thurston's argument. Take a maximal set of points of M with the property that no two of the points have distance $\leq \varepsilon/2$; maximality insures that the $\varepsilon/2$-balls cover the thick part. The $\varepsilon/4$ balls about such points are mutually disjoint. The total volume of the $\varepsilon/4$-balls cannot exceed V so there are a finite number. The combinatorial pattern of intersections of the $\varepsilon/2$-balls determines the homeomorphism type of M; there are only a finite number of possibilities.

Unfortunately, as pointed out in [Benedetti and Petronio 1992, pp. 195-6] it is possible that a ε-tube may bore though an $\varepsilon/2$-ball, leaving one or more worm holes. This increases the possibilities for the topological type of M, beyond what is accounted for above. For this reason subsequent authors have to find lengthy alternate treatments to avoid this difficulty among others; see [Petronio and Porti 2000].

The finiteness of topological types is coupled with the fact that two manifolds of finite volume which have homeomorphic ε-thick parts can be obtained from one another by Dehn surgery. As a consequence, given V, all manifolds of volume $\leq V$ are obtained by Dehn surgery on the cusp tori of a finite number of manifolds.

To analyze (iii), suppose a sequence of volumes is strictly decreasing. After passing to another subsequence if necessary, we may assume the groups G_n have a geometric limit H. Then $\mathrm{Vol}\mathcal{M}(H) = \lim \mathrm{Vol}\mathcal{M}(G_n)$. By Theorem 4.10.1 the geometric limit has at least one more rank two cusp than its close approximates. By Theorem 4.6.1, the close approximates arise from Dehn surgery on the rank two cusps of the geometric limit. By item (ii), the close approximates have lower volume.

On the one hand the volume of $\mathcal{M}(H)$ is greater than the volume of its close approximates, and on the other, the volume of its approximates is strictly decreasing. This contradiction proves that the volumes of the sequence must stabilize at a certain point, as claimed.

Item (iv) holds because there are only a finite number of homeomorphism types of the solid cusp tori complements of elements of $\mathfrak{M}_V$. □

The well ordering

Suppose there is at least one noncompact manifold of volume V. The set of manifolds with volume V serves as the “mothers” of the manifolds $\{\mathcal{M}(G)\}$ of volume $< V$ with the following property. There are a finite number of mutually disjoint nontrivial simple loops, which one can think of as forming a link $L = \bigcup_i \gamma_i \subset \mathcal{M}(G)$, for which $\mathcal{M}(G) \setminus \bigcup_i \gamma_i$ is homeomorphic to a mother $\mathcal{M}(H)$. Each mother $\mathcal{M}(H)$ is the geometric limit of the manifolds $\mathcal{M}(G_n)$ obtained by Dehn surgeries on its cusp tori.

If we start with the set of noncompact manifolds of lowest possible volume V, then their set of children comprise all closed hyperbolic manifolds of volume $< V$.

Theorem 4.10.1 leads to the conclusion that the set of volumes is well ordered (every subset has a least element):

$$v_1 < v_2 < \cdots \longrightarrow v_\omega \ < v_{\omega+1} < v_{\omega+2} < \cdots \longrightarrow v_{2\omega} < \cdots \longrightarrow v_{\omega^2} \ < \cdots .$$

Here v_ω is the lowest volume for 1-cusped manifolds: v_1 is the lowest volume for closed hyperbolic manifolds, v_2 the second lowest, and so on, so that v_ω is the least accumulation point of volumes of closed manifolds obtained by Dehn surgery on the least volume 1-cusped manifolds. Here ω is the ordinal of the positive integers. Then $v_{2\omega}$ is the next lowest volume of 1-cusped manifolds and is the accumulation point of volumes $v_{\omega+1}, v_{\omega+2}, \ldots$ obtained by Dehn surgery on these. And so on until reaching the first accumulation point v_{ω^2} of volumes $v_{k\omega}$ of 1-cusped manifolds; v_{ω^2} is the lowest volume for 2-cusped manifolds. This spawns the volume sequence $v_{2\omega^2}, v_{3\omega^2}, \ldots$ of 2-cusped manifolds which in turn accumulates at the least volume v_{ω^3} for 3-cusped manifolds. Here ω stands for the cardinal number of the integers.

The index t of a general element v_t of the volume sequence is an ordinal number of the form

$$m_n\omega^n + m_{n-1}\omega^{n-1} + \cdots + m_0,$$

where m_j is a nonnegative integer. For example the index $2\omega^2 + 4\omega + 6$ corresponds to the volume of a manifold obtained first by Dehn surgery on the 2nd lowest volume 2-cusped manifold resulting in a 1-cusped manifold of the 4th lowest volume followed by surgery resulting in a closed manifold with 6th lowest volume.

Thus the set of volumes form successive intervals on $\mathbb{R}$ of the form $[0, \omega]$, $[0, \omega^2]$, $\ldots$, $[0, \omega^\omega)$. The order type of the set of all volumes is the ordinal ω^ω.

In [Cao and Meyerhoff 2001] it is shown that $v_\omega = 2v \cong 2.03$, where v is the volume of the regular ideal tetrahedron, and that among the cusped manifolds, only the complement of the figure-8 knot Figure 3.8 (p. 164) and its sibling in $\mathbb{S}^3$ achieve it (see Exercise 3-5). Among all (orientable) manifolds, the minimum volume can be attained only by a closed manifold. It is conjectured that the minimum is $v_1 = 0.9427\ldots$, that value being attained by the Weeks manifold obtained by $(5, 1)$, $(5, 2)$ Dehn surgery on the two components of the Whitehead link. Several people, such as I. Agol [2004] are currently working to find the minimum volume manifold; the best result to date is that of A. Przeworski: $v_1 > 0.3325$.

For a report on the cusped hyperbolic manifolds composed of at most seven ideal tetrahedra and their Dehn surgery daughters, see [Callahan et al. 1999].

The discovery of the minimal volume orientable *orbifold* has recently been announced by Marshall and Martin [$\geq$ 2007]. It comes from an order two extension of the orientation preserving subgroup of the reflection group of the following hyperbolic tetrahedron: Two faces form a $\pi/5$-dihedral angle and each of these faces form a $\pi/3$-dihedral angle with another; the remaining three dihedral angles are $\pi/2$. The minimum volume orbifold is uniquely determined and has volume $0.03905\ldots$. Its discovery allows the investigation of maximal automorphism groups of closed manifolds; see [Conder et al. 2005]. Earlier Meyerhoff [1987] had shown that the group H of orientation preserving symmetries of the tessellation of $\mathbb{H}^3$ by regular ideal tetrahedra gives the smallest volume orientable orbifold with one cusp.

The well ordering of volumes of finite volume hyperbolic orbifolds is shown in [Dunbar and Meyerhoff 1994].

Volumes of higher-dimensional manifolds

It is interesting to contrast the situation of 3-dimensional finite volume manifolds with other dimensions. The areas of finite area 2-dimensional hyperbolic manifolds are integral multiples of 2π (Exercise 3-1). For a finite volume even dimensional hyperbolic manifold M^{2n}, the formula also comes from the Gauss-Bonnet formula, for example see [Kellerhals and Zehrt 2001].

$$\mathrm{Vol}(M^{2n}) = (-1)^n \frac{V_{2n}}{2}\chi(M^{2n}),$$

where V_{2n} is the surface area of the unit $(2n-1)$-dimensional sphere * in $\mathbb{R}^{2n}$ and χ denotes the Euler characteristic. The formula holds for orientable and nonorientable manifolds; in the former case the Euler characteristic is even for a closed manifold. The paper [Ratcliffe and Tschantz 2000] explicitly constructs the finite volume cusped (noncompact) hyperbolic 4-manifolds. Exactly 1171 of them have the minimum volume.

For odd-dimensional finite volume hyperbolic manifolds, the Euler characteristic is zero. However for *cusped* manifolds, using ball packing methods, good lower bounds can be found [Adams 1987; Kellerhals 1998].

It is known that in every dimension $n \geq 4$ the number of nonisometric manifolds with volume less than any prescribed number is finite [Wang 1972]; thus the set of volumes is a discrete set on $\mathbb{R}$. Furthermore, the number $N(V)$ of nonisometric manifolds of volume $\leq V$ grows to $+\infty$ with V; in fact, it is shown in [Burger et al. 2002] that there are constants $a = a(n) > 0$, $b = b(n) > 0$ such that for all large V,

$$e^{aV \log V} \leq N(V) \leq e^{bV \log V}.$$

4.12 Exercises and explorations

4-1. Prove that if $\langle U, V\rangle$ is discrete and nonelementary, the subgroup $\langle U, [U, V]\rangle$ is also nonelementary provided U is not elliptic of order ≤ 60.

4-2. Suppose that the sequence of kleinian groups $\{G_k\}$ converges polyhedrally to a geometrically finite group H. Prove that there is a homomorphism ψ_k of H into G_k, for all large k, such that $\lim \psi_k(h) = h,\ h \in H$.

4-3. [Mumford 1971] Prove that the collection of all closed Riemann surfaces (compact surfaces without boundary) of genus $g \geq 2$ that have the property that the length of any closed geodesic exceeds some $\varepsilon > 0$ is compact: Every infinite sequence of such surfaces, or infinite sequence of normalized fuchsian covering groups, has a geometrically convergent subsequence to a group which represents a surface of the same type.

4-4. In contrast to the example of Section 4.10, verify the following claim. A sequence of cyclic loxodromic groups $\{\langle S_n\rangle\}$ with real traces which converges algebraically to the cyclic parabolic group $\langle S\rangle$ also converges to it geometrically.

Show that the conclusion remains the same if the hypothesis is weakened to the assumption that there exists $\delta > 0$ such that for all indices,

$$-\frac{\pi}{2} + \delta \leq \arg(\operatorname{tr} S_n) \leq \frac{\pi}{2} - \delta.$$

Looking at the quotients, the sequence of solid tori converge geometrically to a solid cusp tube.

* The volume V_k of $\mathbb{S}^{k-1} \subset \mathbb{R}^k$ is $2\pi^{n/2}\Gamma(n/2)$.

4-5. We will follow [Thurston 1979, Lemma 5.6.1] in outlining the sufficiency of Equation 4.2 for the extension of ψ from a homomorphism of $\pi_1(M')$ to one of $\pi_1(M)$ (compare Lemma 4.11.1). The proof proceeds by considering each boundary torus separately. For simplicity of notation, we may therefore assume there is only one.

We have chosen generators of the genus two surface S so that $[\gamma,\delta][\alpha,\beta]=\mathrm{id}$. We are assuming that $\mathrm{tr}(\psi([\alpha,\beta])=2$, and $\mathrm{tr}\psi(\beta)=2$. We are also assuming that $\psi(\langle\gamma,\delta\rangle)\neq\mathrm{id}$, $\psi(\alpha)$ is loxodromic (does elliptic and parabolic also work?), and $\psi(\langle\alpha,\gamma,\delta\rangle)$ is nonelementary. According to Lemma 1.5.2, $\psi(\alpha)$ and $\psi(\beta)$ have a common fixed point, say ∞.

Take $\psi(\alpha)=\left(\begin{smallmatrix}\lambda & 0\\ 0 & \lambda^{-1}\end{smallmatrix}\right)$. Then $\psi(\beta)=\left(\begin{smallmatrix}a & b\\ 0 & a^{-1}\end{smallmatrix}\right)$. If $\psi([\alpha,\beta])$ is not the identity, then since $\psi([\gamma,\delta])=\psi([\alpha,\beta]^{-1})$, all four of $\psi(\alpha),\psi(\beta),\psi(\gamma),\psi(\delta)$ have ∞ as a fixed point, a contradiction. Finally if $\psi(\beta)\neq\mathrm{id}$, then $\psi(\beta)$ is parabolic. This too is impossible for then it could not commute with $\psi(\alpha)$.

4-6. *Convergence of limit sets.* [McMullen 1996, Prop. 2.4] Suppose $\{G_n\}$ is a sequence of kleinian groups normalized so that the respective convex hulls $\widetilde{\mathcal{C}}(G_n)$ contain a fixed ball about a point $O\in\mathbb{H}^3$ that projects injectively to the quotients. We may then assume that the sequence converges geometrically to a kleinian group H.

A point z lies in $\liminf\Lambda(G_n)$ when every neighborhood U of z contains points of $\Lambda(G_n)$ for all n, with at most a finite number of exceptions. In contrast $z\in\limsup\Lambda(G_n)$ when every neighborhood U of z contains points of infinitely many $\Lambda(G_n)$. The sequence $\{\Lambda(G_n)\}$ converges when the two limits agree.

Because every loxodromic fixed point of H is the limit of loxodromic fixed points of G_n, conclude that $\Lambda(H)\subset\liminf\Lambda(G_n)$.

There is not always equality in the two limits. For example, there is a sequence of fuchsian groups of the first kind whose geometrical limit is just {id} — which has empty limit set. One such is the sequence of level-n subgroups $\{M_n\}$ of the modular group (Exercise 2-9: $M_n=\{g\in\mathrm{Mod}: g\equiv I \bmod n\}$).

But suppose that there exists $\rho<\infty$ such that $\widetilde{\mathrm{Inj}}_n(x)<\rho$ for all $x\in\widetilde{\mathcal{C}}(G_n)$ and all n. Prove that $\lim\Lambda(G_n)=\Lambda(H)$ in the Hausdorff topology.

Hint: For a kleinian group G and $r<\infty$ set $M(r)=\{x\in\mathbb{H}^3:\widetilde{\mathrm{Inj}}(x)\le r\}$. First show that $\overline{M}(r)\cap\mathbb{S}^2\subset\Lambda(G)$, where $\overline{M}(r)$ denotes closure in the spherical metric. To see this descend to $\mathcal{M}(G)$. Suppose that $\mathrm{Inj}(\pi(x))\le r$ for $\pi(x)\in\mathcal{M}(G)$. Then there is a noncontractible closed loop c of length $\le 2r$ through $\pi(x)$. Shrink it so that it either becomes a geodesic $\gamma\subset\mathcal{C}(G)$ of length $\le 2r$ or a simple loop γ on the boundary of a thin part at a finite distance from c. Back upstairs, in the former case, the G-orbit of x accumulates to $\Lambda(G)$. In the latter case, the orbit of x under a cyclic parabolic group accumulates to a parabolic fixed point.

Returning to our geometrically convergent sequence, show that $\{\widetilde{\mathrm{Inj}}_n(x)\}$ for $\{G_n\}$ converges to $\widetilde{\mathrm{Inj}}(x)$ for H, uniformly on compact subsets of $\mathbb{H}^3$. That is, lim sup $M_n(r)\subset M(r)$ with respect to Hausdorff convergence, for any $r>0$. When $r=\rho$, $\widetilde{\mathcal{C}}(G_n)\subset M_n(\rho)$, so $\limsup\widetilde{\mathcal{C}}(G_n)\subset M(\rho)$. Since all rays from $O\in\widetilde{\mathcal{C}}(G_n)$ to points on $\Lambda(G_n)$ lie in $\widetilde{\mathcal{C}}(G_n)$, the limiting rays lie in $M(\rho)$. Conclude that $\limsup \Lambda(G_n) \subset \overline{M}(\rho) \subset \Lambda(H)$.

4-7. *Hyperbolic cone manifolds and orbifolds* We will start with the simplest of examples. Set $\alpha = 2\pi/n$ and consider the action of the elliptic $E_\alpha(z) = e^{i\alpha}z$ on $\mathbb{H}^3$. The quotient $\mathcal{M}(\langle E_\alpha \rangle)$ is an (oriented) orbifold. The *cone angle* at the set Σ of cone points, which is just the projection of the rotation axis, is $2\pi/n$. The elliptic E_α may be conjugated so that as $n \to \infty$, it converges to a parabolic (Exercise 2-4). The quotient $M_{\mathcal{O}} = \mathbb{H}^3 \setminus \text{cone axis}/\langle E_\alpha \rangle$ has a hyperbolic metric but it is not complete. It does have a metric completion which is topologically a ball B and is called the *underlying space* of the orbifold. In B, the projection of the rotation axis is called the *singular locus* $\Sigma_{\mathcal{O}}$.

What about letting α be an irrational multiple of 2π? On the one hand it is perfectly reasonable for a cone to have an irrational cone angle. On the other hand the group is no longer discrete. Still we can use the model $M_{\mathcal{O}}$ for it. Again the underlying space is B and $M_{\mathcal{O}} \cong B \setminus \Sigma_{\mathcal{O}}$. The singular set $\Sigma_{\mathcal{O}}$ has cone angle α. We have no covering map π^{-1} to use, instead we have a more general map called a *developing map* that operates in the same way. The developing map d is a local isometry, unrolling $M_{\mathcal{O}}$ in $\mathbb{H}^3$: Given a closed loop γ encircling $\Sigma_{\mathcal{O}}$, and a point O^* that projects to its initial point, the developing map "lifts" γ, starting at O^* and terminating at $E_\alpha^k(O^*)$ for some k. The developing map is coupled with a homomorphism φ, called the *holonomy map*. The holonomy map sends $\pi_1(M_{\mathcal{O}})$ to the group of Möbius transformations generated in this case by E_α.

A 3-manifold M_C is called a *hyperbolic cone manifold* if there is a link $\Sigma \subset M_C$ (a union of simple loops not isotopic in $M_C \setminus \Sigma$ to a point) with the following properties.

(i) $M_C \setminus \Sigma$ has an incomplete hyperbolic structure.
(ii) Its metric completion in M_C is a singular metric with cone type singularities along Σ.
(iii) To each component σ of Σ corresponds a *cone angle* $0 < \alpha < \infty$. In a thin tubular neighborhood $N_r(\sigma)$ of radius r about σ the metric can be expressed in cylindrical coordinates (see Exercise 8-8) as

$$dr^2 + \sinh^2 r\, d\theta + \cosh^2 r\, dh^2,$$

where H is the distance along σ and θ (mod α) is the angular measure about σ.

In other words there is a developing map $D : M_C \setminus \Sigma \to \mathbb{H}^3$ which is a local isometry that unrolls $M_C \setminus \Sigma$ in $\mathbb{H}^3$. The meridian on $\partial N_r(\sigma)$ lifts to an elliptic transformation E. Any longitude lifts to a loxodromic transformation with the same axis as E. The lift D^* to the universal cover of $M_C \setminus \Sigma$ induces a homomorphism from $\pi_1(M_C \setminus \Sigma)$ to $\mathrm{PSL}(2, \mathbb{C})$.

A cone manifold with rational cone angles $2\pi/n$ is an orbifold—but in general the singular set of an orbifold is not a link (Proposition 2.5.2). The limiting cases of cone angle zero corresponds to a rank two cusp while cone angle 2π signifies that there is no singularity at σ.

Assume that $M_C \setminus \Sigma$ has finite volume and the cone angles along Σ are at most 2π. The fundamental local rigidity theorem of Hodgson and Kerckhoff [1998] asserts that the set of cone angles provides a local parameterization of the hyperbolic cone structures near the given one. In particular there are no infinitesimal deformations of the hyperbolic structure that keep the cone angles fixed. In fact $M_C \setminus \Sigma$ has a complete hyperbolic structure of finite volume [Kojima 1996]. If in addition the angles are at most π, there is a continuous, angle-decreasing family of deformations converging to the complete structure [Kojima 1998]. He also showed that if cone manifolds $M_C \setminus \Sigma$, $M'_C \setminus \Sigma'$ are homeomorphic with corresponding cone angles the same (and all at most π), the two manifolds are isometric.

One natural source of cone manifolds, used for example in Theorem 3.11.3, is the reflection in its boundary of a convex core for which the bending locus consists of simple geodesics. The bending lines become the singular locus and the cone angles are twice the interior bending angles. In this case the cone angles are less than 2π so the Hodgson-Kerckhoff deformation theory is operative.

For a recent study of cone manifolds with singular locus the union of circles and trivalent graphs (as with orbifolds), with cone angles $\leq \pi$, see [Weiss 2002].

For more discussion see Exercise 6-3.

4-8. *Geometric limits of fuchsian groups.* In the surface $R = \mathbb{D}/G$ let γ be a simple loop cutting R in two components R_1, R_2. Let $R_1^* \subset \mathbb{D}$ be a lift of R_1 and let R_2^* be a lift (a component of the preimage) of R_2 such that R_2^* is adjacent to R_1^* along a lift γ^* of γ. Let $G_i \subset G$ denote the stabilizer of R_i, $i = 1, 2$ and g^* a generator of the stabilizer of γ^*. Show that $G = \langle G_1, G_2 \rangle$ (in the language of combinatorial group theory, G is the free product of G_1 and G_2 with amalgamation over the cyclic subgroup that stabilizes γ^*). There is a general way to find a sequence of isomorphisms $\theta_n : G \to H_n$ to other fuchsian groups $\{H_n\}$ such that $\lim \theta_n(g^*)$ is parabolic and $\lim \theta_n(g)$ exists as a Möbius transformation for all $g \in G_1$. On the other hand for $g \in G_2$ with $g(\gamma^*) \neq \gamma^*$, no subsequence of $\theta_n(g)$ converges to a Möbius transformation, that is, the regions $\{R_{2,n}^*\}$ shrink to the fixed point of $\lim \theta_n(g^*)$.

This can be done as follows. First note that a sequence of increasingly thick annuli can always be normalized to converge to a once punctured disk, for example the sequence $\{1/n < |z| < 1\}$. Cut R along γ and sew back in increasingly thick annuli thereby getting new surfaces $R^n = R_{1,n} \cup R_{2,n}$. Apply the Uniformization Theorem to get a new fuchsian group H_n and an isomorphism $\theta_n : G \to H_n$. We are free to replace H_n by a conjugate; do so that for all n the lift $R_{1,n}^*$ contains a fixed small disk about 0 that embeds in the quotient. The element $\theta_n(g^*) \in H_n$ and all its conjugates with respect to $H_{1,n}$ converge to parabolic transformations. The groups $\theta_n(G_1)$ converge algebraically to a fuchsian group representing a surface homeomorphic to R_1 but with a puncture in place of the boundary component γ. The subgroup of H_n that preserves any lift of R_2^n, or any lift of R_1^n other than R_1^{n*}, degenerates.

a puncture in place of the boundary component γ. The subgroup of H_n that preserves any lift of R_2^n, or any lift of R_1^n other than R_1^{n*}, degenerates.

The geometric limit of $\{H_n\}$ is the algebraic limit of $\{H_{1,n}\}$. Note the role of the choice of basepoint as the focus of conjugation. One could have chosen it so that the convergent sequence was instead $\{H_{2,n}\}$. This is only one example in a complete description of all geometric limits of sequences of fuchsian groups. This process was first described by Bill Harvey [1977]. Can you formulate a general theorem describing all possible geometric limits?

Now consider the kleinian case. On the boundary of a geometrically finite $\mathcal{M}(G)$, suppose $\mathcal{S}$ is a set of mutually disjoint, noncompressing simple loops with the property that no two are freely homotopic within $\mathcal{M}(G)$.

Deformation theory allows the deformation of the group so that all the curves of the set $\mathcal{S}$ become "pinched"; the corresponding elements of G become parabolic. The example of the two surfaces resulting from a fuchsian group shows why the homotopy property is a necessary condition to carry this out. There results a new geometrically finite group H in various interesting combinatorial arrangements depending on $\mathcal{S}$. Show that a genus two Schottky group can be so pinched as to become a pair of once punctured tori. See the Pinching Theorem (page 286) and Exercise 5-3. Each once punctured torus can be pinched at most once again in countably many ways so as to become a thrice punctured sphere. For an elementary and detailed presentation of the two generator Schottky case, see [Mumford et al. 2002].

4-9. *Isomorphisms determining homeomorphisms.* Suppose G is a geometrically finite group without parabolics and $\mathcal{M}(G)$ is acylindrical. If Ω_1, Ω_2 are distinct components of $\Omega(G)$ and $G_i = \mathrm{Stab}(\Omega_i)$, prove that $\Lambda(G_1)\cap\Lambda(G_2)=\varnothing$ (see [Matsuzaki and Taniguchi 1998, §3.2.1, Theorem 3.29]). Conclude that if $\varphi : G \to H$ is an isomorphism to another geometrically finite group without parabolics then $\mathcal{M}(H)$ is also acylindrical and there is a bijection between the components of $\Omega(G)$ and $\Omega(H)$ (see Exercise 3-29).

Hint: Assume to the contrary that $\Lambda(G_1)\cap\Lambda(G_2)\neq\varnothing$. Take a ray ℓ in the convex hull $\hat{\mathcal{C}}(G_2)$ of $\Lambda(G_2)$ ending at $\zeta \in \Lambda(G_1)\cap\Lambda(G_2)$. Given $O \in \mathbb{H}^3$ there exists $\{g_n \in G_1\}$ such that $\{g_n(O)\}$ lies in a conical neighborhood of ℓ. Project the sequence into $\mathcal{M}(G_2)$. It is contained in a finite distance neighborhood of the convex core $\mathcal{C}(G_2)$ of $\mathcal{M}(G_2)$. Therefore there is a sequence $h_n \in G_2$ such that (for a subsequence) $\lim h_n g_n(O) = O' \in \mathbb{H}^3$. Discreteness requires that $h_n g_n = \mathrm{id}$ for all large indices. Therefore $h_n \in G_1 \cap G_2$, a contradiction.

4-10. *Alternate definitions of geometric convergence.* A quantitative version of the definition of Section 4.2 goes as follows. Suppose we have a sequence $\{G_n\}$ and a point $O \in \mathbb{H}^3$ such that the corresponding fundamental polyhedra $\{\mathcal{P}_n\}$ with origins at O all contain small ball B_ϵ about O with $g_n(B_\epsilon)\cap B_\epsilon=\varnothing$ for all $g_n \neq \mathrm{id} \in G_n$ and all indices n.

The sequence converges geometrically to the group H if and only if the following holds: There exists a sequence of K_n-bilipschitz maps $F_n : \mathbb{H}^3 \to \mathbb{H}^3$, $F_n(O) = O$, with $\lim K_n = 1$, which have the following two properties. (i) On every compact subset of $\mathbb{H}^3$, $\lim F_n(x) = x$. (ii) Choose a sequence $r_n > 0$, $\lim r_n = \infty$. Set $M_n = (\mathcal{P}_n \cap B_{r_n})/G_n$, where B_{r_n} is the ball of radius r_n about O. Then F_n projects to a K_n-bilipschitz map f_n of M_n into $\mathbb{H}^3/H$, see [Canary and Minsky 1996, Lem. 3.1].

From a more general perspective, Gromov presented the following definition of geometric convergence [Gromov 1981b].

We are given two metric spaces X, Y and a map $f : X \to Y$. Using the respective metrics set

$$L(f) = \sup_{x_1 \neq x_2 \in X} \left| \log \frac{d(x_1, x_2)}{d(f(x_1), f(x_2))} \right| .$$

A sequence of metric spaces $\{(X_n; O_n)\}$ with basepoints O_n is said to *converge* to the metric space with basepoint $(Y; O)$ if the following holds. Given any $r > 0$ and $\varepsilon > 0$, there exists N such that for each $n \geq N$, there exists a map f_n from the radius-r ball $B_r(O_n) \subset X_n$ into Y such that

(i) $f_n(O_n) = O$,
(ii) $f_n(B_r(O_n)) \subset Y$ contains the ball $B_{r-\varepsilon}(O) \subset Y$, and
(iii) $L(f_n) \leq \varepsilon$, computed on $B_r(O_n)$.

For an application to our situation, set $X_n = \mathcal{P}_O(G_n) \cap B_r(O)/G_n$ and $Y = \mathbb{H}^3/H$.

4-11. $\mathbb{R}$-*trees.* Another way of representing the degeneration of hyperbolic manifolds is by $\mathbb{R}$-trees. This point of view was pioneered by Morgan and Shalen [1984; 1988a; 1988b]. (See also references in [Bestvina 1988; Ohshika 1998b; Otal 1996]). Here we will define $\mathbb{R}$-trees and in the next exercise show how they arise in degenerations of kleinian groups.

A metric space $\mathcal{T} = (X, d)$ is called a *real tree* or $\mathbb{R}$-*tree* if there is a *unique* arc (up to parameter change) connecting any two points $x, y \in X$ and that arc has length $d(x, y)$.

Thus if $[a, b, c] \subset X$ is a triangle, each side is contained in the union of the other two. This is called the *tripod property*. The center of the tripod is the unique point $[a, b] \cap [b, c] \cap [c, a]$. Conversely suppose a metric space (Y, d) has the properties that (i) there is an arc of length $d(x, y)$ between any two points $x, y \in X$, and (ii) the tripod property holds for any geodesic triangle. Then there is a unique arc between any two points, and (Y, d) is an $\mathbb{R}$-tree.

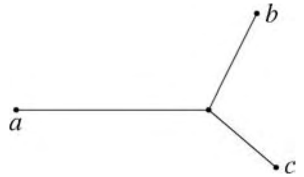


Fig. 4.2. A tripod.

For example, let $X \subset \mathbb{C}$ be the union of three rays from the origin and let d be the metric on X induced from $\mathbb{C}$. Then X is an $\mathbb{R}$-tree.

Another example comes from taking the "dual graph" of a measured lamination (Λ, μ) in $\mathbb{H}^2$, the lift of one on a closed surface S of genus exceeding one [Otal 1996, §2.3]. For simplicity we will assume that Λ is minimal without closed leaves.

The points of $\mathcal{T}$ will be of two types: (i) the closure of a component (gap) of $\mathbb{H}^2 \setminus \Lambda$, and (ii) a leaf $\lambda \subset \Lambda$ that is not in such a closure. Define the distance between the points as follows. Suppose $x, y \in \mathcal{T}$ are points that correspond to two gaps. Choose points, also denoted x, y, in each gap and consider the geodesic segment $[x, y]$ between them. Then the positive number $\mu[x, y]$ is the transverse measure of the segment. "Integration" of this measure determines a distance between the closed subsets that intersect $[x, y]$. This definition is independent of the choice of x, y in their gaps. Given any two points of $\mathcal{T}$, there exist gaps that separate them. Thus the distance $d(\cdot, \cdot)$ can be defined between any pair of points of $\mathcal{T}$.

With this distance, $\mathcal{T}$ is an $\mathbb{R}$-tree. If (Λ, μ) is invariant under the action of a fuchsian group G, the action of G determines a fixed point free isometry of $\mathcal{T}$. The action is also minimal. Conversely, we have the following basic theorem of Skóra (see [Otal 1996, Theorem 2.3.5]):

Theorem 4.12.1. *Suppose $G \times Y \to Y$ is a nontrivial isometric, minimal action of the fuchsian group G on the $\mathbb{R}$-tree Y. Assume that every subgroup of G that fixes an arc of Y has a finite index abelian subgroup, and that the action of parabolic elements of G have translation distance zero in Y. Then the action of G on Y is isometric to the action of G on the tree $\mathcal{T}$ determined by a measured geodesic lamination in $\mathbb{H}^2$.*

4-12. *$\mathbb{R}$-trees and the degeneration of manifolds.* Because of its significance in the general theory of group deformations, we will provide a somewhat lengthy outline of the approach in [Bestvina 1988] to showing how degeneration results in an $\mathbb{R}$-tree. This approach has proved useful in establishing hyperbolization for fibered 3-manifolds [Otal 1996; Kapovich 2001].

Suppose G is a (nonelementary, finitely generated) kleinian group. Fix a set of generators $\{g_1, g_2, \ldots, g_r\}$ of G. Let $\{\theta_n : G \to G_n\}$ be a sequence of isomorphisms to discrete groups. Renormalize if necessary so that a given basepoint point $O \in \mathbb{H}^3$ is moved least by these generators in the sense that

$$d_n = d(G_n) = \max_{1 \le i \le r}\{d(O, \theta_n(g_i)(O))\} \le \max_{1 \le i \le r}\{d(\vec{x}, \theta_n(g_i)(\vec{x}))\}$$

for all $\vec{x} \in \mathbb{H}^3$. Here $d(\cdot, \cdot)$ denotes hyperbolic distance. Show that if the sequence $\{d_n\}$ is uniformly bounded, a subsequence can be chosen so that $\{\theta_k\}$ converges to an isomorphism θ.

Since we want to study degenerations, assume that $\lim d_n = \infty$.

Let $\mathcal{W}^k$ denote the set of words in the given generators of G which have length $\le k$. Let $\mathcal{F}_n^k$ denote the convex hull in $\mathbb{H}^3$ of the point set $\{\theta_n(g)(O) : g \in \mathcal{W}^k\}$.

Now rescale $\mathcal{F}_n^k$: let $\mathcal{X}_n^k$ denote the abstract metric space whose set of points is $\mathcal{F}_n^k$ but the distance between two points is rescaled as $\rho_n(x, y) = d(x, y)/d_n$, for $x, y \in \mathcal{F}_n^k$. Thus $\rho_n(x, y) \le 2$.

A sequence $\{Y_n\}$ of compact, connected metric spaces is said to converge in the *Gromov sense* to the metric space Y if the following holds. There is a compact metric space Z and isometric embeddings $Y_n \hookrightarrow Z$ and $Y \hookrightarrow Z$ so that as subsets of Z, $\{Y_n\}$ converges to Y in the Hausdorff topology. That is, in terms of the images in Z, given any $\varepsilon > 0$ there exists $n(\varepsilon)$ so that when $n > n(\varepsilon)$, Y is contained in the ε-neighborhood of Y_n and Y_n is contained in the ε-neighborhood of Y. Gromov showed that the limit is uniquely determined up to isometry.

Gromov [1981a, §7] proved that a necessary condition on the metric spaces for there to exist a Gromov convergent subsequence is that to every $\varepsilon > 0$, there is an integer $N(\varepsilon)$ such that every Y_n can be covered by $N(\varepsilon)$ ε-balls.

It is interesting to see Gromov's construction of Z. Set $\varepsilon_i = 2^{-i}$. For each i there exists an integer N_i so that every Y_n is covered by N_i ε_i-balls. For each i, introduce the i-tuple of integers $A_i = \{(n_1, n_2, \ldots, n_i) : 1 \leq n_j \leq N_j,\ 1 \leq j \leq i\}$. For fixed n define inductively a sequence of maps of $A_1, A_2, \ldots, A_k, \ldots$ into Y_n as follows:

(1) Cover Y_n by N_1 ε_1-balls; choose any one-to-one map I_n^1 from the set of N_1-integers $A_1 = \{n_1\}$ to the N_1 centers of the ε_1-balls.
(2) Cover each ε_1-ball from the N_2 ε_2-balls; I_n^2 is the map of the set A_2 to the set of centers of the ε_2-balls such that (n_1, n_2) goes to the center of an ε_2-ball used to cover the ε_1-ball centered at $I_n^1(n_1)$.
(3) Cover each ε_2-ball from the N_3 ε_3-balls; I_n^3 is the map of A_3 onto the centers of the ε_3-balls such that (n_1, n_2, n_3) goes to the center of a ball in the cover of the ε_2-ball centered at $I_n^2(n_1, n_2)$. And so on.

Set $A = \bigcup_1^\infty A_k$ and let $I_n : A \to Y_n$ denote the corresponding map. Consider the metric space of maps $B = \{f : A \to \mathbb{R} : \text{f is bounded}\}$ with metric determined by the norm $\|f\| = \sup_{a \in A} |f(a)|$. Let $Z \subset B$ be the compact metric subspace of those functions satisfying

$$\begin{array}{llll} \text{if} & a \in A_1, & \text{then} & 0 \leq f(a) \leq \sup_n \{\mathrm{Diam} Y_n\}; \\ \text{if} & a \in A_k,\ k > 1, & \text{then} & |f(a) - f(p_{k-1}(a))| \leq 2\varepsilon_{k-1}, \end{array}$$

where $p_k : A_{k+1} \to A_k$ is the natural projection. The construction is such that for each $a \in A_k$, $I_n^k(a)$ is contained in the $2\varepsilon_{k-1}$-ball centered at $I_n^{k-1}(p_{k-1}(a))$.

Define $h_n : Y_n \to B$ by

$$(h_n(x))(a) = \mathrm{dist}(x, I_n(a)), \quad x \in Y_n,\ a \in A;$$

here "dist" is the metric in Y_n. Verify that the image of h_n is contained in Z and that h_n gives an isometric embedding of Y_n.

Finally, the space of all compact subsets of a compact set is itself compact with respect to the Hausdorff topology, so there is a convergent subsequence of $\{Y_n\}$.

Applying Gromov's construction to our situation, we get:

Theorem 4.12.2. *For each k, there is a subsequence of $\{\mathcal{X}_n^k\}$ which in the Gromov sense converges to a compact metric space $\boldsymbol{T}^k$.*

Proof. Given $\varepsilon > 0$ we will verify that there exists $N(\varepsilon)$ such that each $\mathcal{X}_n^k$ can be covered by $N(\varepsilon)$ ε-balls, or equivalently, $\mathcal{F}_n^k$ can be covered by $N(\varepsilon)$ $d_n\varepsilon$-balls. If $\mathcal{W}^k$ has $W(k)$ elements, then $\mathcal{F}_n^k$ has $W(k)$ vertices and $\frac{1}{2}W(k)(W(k)-1)$ diagonals (geodesic segments between distinct vertices). The length of each diagonal cannot exceed $2kd_n$. Cover the diagonals by εd_n-balls centered at points on the diagonals spaced at distance $\leq \varepsilon d_n$. This requires at most

$$\frac{W(k)(W(k)-1)}{2}\left(\left[\frac{2kd_n}{\varepsilon d_n}\right]+1\right) = \frac{W(k)(W(k)-1)}{2}\left(\left[\frac{2k}{\varepsilon}\right]+1\right)$$

balls.

Now, by Exercise 1-24, there exists a constant $C > 0$ such that for each $\vec{x}$ in any triangle or in any tetrahedron in $\mathbb{H}^3$, the shortest distance of $\vec{x}$ to the edges does not exceed C. Therefore the collection of balls covers $\mathcal{F}_n^k$ if $\varepsilon d_n \geq 2C$, since each point of $\mathcal{F}_n^k$ lies in a tetrahedron whose vertices are among the vertices of $\mathcal{F}_n^k$.

Therefore, after passing to a subsequence if necessary and taking a diagonal subsequence, we can arrange matters so that (1) $\mathcal{X}_n^k \to \boldsymbol{T}^k$ in the Gromov sense, (2) $\cdots \subset \boldsymbol{T}^{k-1} \subset \boldsymbol{T}^k \subset \boldsymbol{T}^{k+1} \subset \cdots$, and (3) for $g \in \mathcal{W}^k$, $\lim_{n\to\infty} \phi_n(g) = x^*(g) \in \boldsymbol{T}^k$ exists. □

In short, the rescaled diagonals of $\mathcal{F}_n^k$ converge to segments or points in an ambient compact metric space and the rescaled convex hulls collapse upon them. More precisely:

Theorem 4.12.3. *$\boldsymbol{T}^k$ is a finite $\mathbb{R}$-tree, meaning that:*

(1) *Any two points of $\boldsymbol{T}^k$ can be joined by a segment, that is, a subspace isometric to a closed interval or a point. The segments can be chosen with endpoints in the set $\{x^*(g) : g \in \mathcal{W}^k\}$.*
(2) *The intersection of any two nondisjoint segments in $\boldsymbol{T}^k$ is a segment or point.*
(3) *The union of two segments with a common endpoint is again a segment.*

That $\boldsymbol{T}^k$ is finite means it is the union of finitely many segments.

The diagonals in $\mathcal{F}_n^k$ give rise to segments in $\mathcal{X}_n^k$ that converge to segments or points in $\boldsymbol{T}^k$; the endpoints converge to points $x^*(e) \in \boldsymbol{T}^k$, $e \in \mathcal{W}^k$. We claim that the limiting segments and points cover $\boldsymbol{T}^k$. If not, for some $x \in \boldsymbol{T}^k$ and some ε, the ε-ball about x does not intersect any of the limiting segments or points. Suppose $x_n \in \mathcal{X}_n^k$ converges to x. For large n the $\mathcal{X}_n^k$-distance between x_n and the rescaled diagonals exceeds $> \varepsilon/2$. In $\mathcal{F}_n^k$, this distance exceeds $\varepsilon d_n/2$. Since we have taken ε so that $\varepsilon d_n \geq 2C$, the corresponding point $x_n \in \mathcal{F}_n^k$ lies in an εd_n-ball centered at a point on the diagonal, a contradiction.

For the remainder of the proof see [Bestvina 1988].

Now set $\boldsymbol{T} = \bigcup_k \boldsymbol{T}^k$. Then $\boldsymbol{T}$ is a metric space with the three properties of the theorem above. Therefore it is an $\mathbb{R}$-tree. The theory continues by studying the action of G on $\boldsymbol{T}$. The basic result is that it acts by isometries so that $gz(h) = z(gh)$ for $z \in \boldsymbol{T}$ and $g, h \in G$, where $z(g) = \lim \theta_n(g)(z_n)$ and $z = \lim z_n$. Deeper study leads to a compactification of the space of algebraic limits and a proof of Thurston's compactness theorem (page 206).

4-13. *The isoparametric inequality for* $\mathbb{H}^3$. This states [Chavel 1993, §6.4] that if $X \subset \mathbb{H}^3$ is a compact set with piecewise smooth boundary and volume $\mathrm{Vol}(X)$, then the surface area of ∂X strictly exceeds the surface area of the hyperbolic ball of volume $\mathrm{Vol}(X)$, unless X itself is a ball. Referring back to the formulas on page 16, we find that the area of ∂X exceeds $2\mathrm{Vol}(X)$.

A variation of an argument of [Cooper 1999] yields the following interesting fact.

Suppose $\mathcal{M}(G)$ *is a closed manifold. There is a presentation of* G *consisting of* m *generators and* $\frac{2}{5}(m+1)$ *relations for which* $\mathrm{Vol}(\mathcal{M}(G)) < \frac{2}{5}\pi(m+1)$.

Proof. Choose a generic polyhedron $\mathcal{P}_{\mathcal{O}}$ such that each vertex is shared by three edges and each edge relation has length three, that is, three polyhedra in the orbit share the edge.

Construct the dual graph Γ to the orbit of $\mathcal{P}_{\mathcal{O}}$. Recall this is done by connecting the center $\mathcal{O}$ to the centers of the polyhedra of the orbit that share a face with $\mathcal{P}_{\mathcal{O}}$, and so on. The vertices of Γ are the points of the orbit $G(\mathcal{O})$.

Each piecewise geodesic loop in Γ that surrounds a single edge is a geodesic triangle transverse to faces of the orbit of $\mathcal{P}_{\mathcal{O}}$. If two edges e, e' of $\mathcal{P}_{\mathcal{O}}$ are related $e' = g(e)$ for $g \in G$ then the triangle about e' is the g-image of the triangle about e.

The number of triangles in Γ with vertex $\mathcal{O}$ is equal to the number s of edges of $\mathcal{P}_{\mathcal{O}}$. The number of triangles which are inequivalent under G is $s/3$. Corresponding to each vertex of $\mathcal{P}_{\mathcal{O}}$ is a cone in $\mathcal{P}_{\mathcal{O}}$ with vertex $\mathcal{O}$ and with its three sides contained in the geodesic triangles.

The union of all these triangles $\{\Delta\}$ separates $\mathbb{H}^3$ into simply connected polyhedra $\{X\}$ each of which contains exactly one vertex in the orbit of the vertices of $\mathcal{P}_{\mathcal{O}}$. Each component X projects injectively into $\mathcal{M}(G)$.

Since $\mathcal{P}_{\mathcal{O}}$ is a compact convex polyhedron, its boundary is a topological sphere. Euler's formula for the boundary is $E-F+V=2$. The number of vertices V is related to the number of edges E by $V = 2E/3$. The face pairing transformations generate the group and there are $F/2 = m$ of these. Therefore $E = 6(m+1)/5$. There are E edge relations, but the edge cycles all have length three leaving $E/3 = 2(m+1)/5$ independent relations.

Now we bring in the fact that the surface area of each component X exceeds twice its volume. Each geodesic triangle lies in the boundary of exactly two polyhedra X. Down in $\mathcal{M}(G)$ count the distinct projections $\{\pi(X)\}$, which fill up $\mathcal{M}(G)$. There are $N = E/3$ distinct triangles $\{\pi(\Delta)\}$. The sum of the area of the distinct $\{\pi(\Delta)\}$ exceeds the volume of $\mathcal{M}(G)$, since each $\pi(\Delta)$ borders two of the components $\{X\}$ and is counted twice when totaling up the volume. The sum of the areas of the distinct triangles is bounded above by πN, where $N = \frac{2}{5}(m+1)$ is the number of independent edge relations of $\mathcal{P}_0$. $\square$

4-14. *The Gromov norm.* Let $\mathcal{M}$ be a closed hyperbolic 3-manifold. Gromov showed that its volume can be obtained as the limit of a process of approximation by 3-chains:

We consider all continuous maps $\{\sigma\}$ of a standard euclidean regular tetrahedron Δ into $\mathcal{M}$. The theory allows us to assume that $\sigma(\Delta)$ is a hyperbolic tetrahedron. These maps are a basis of the vector space of singular real chain complexes $\{\sum a_i\sigma_i\}$ in $\mathcal{M}$, $a_i \in \mathbb{R}$. The most natural source of 3-chains are triangulations of $\mathcal{M}$ by tetrahedra (these are easily constructed from a fundamental polyhedron).

Recall from Exercise 1-23 that among all tetrahedra, the regular ideal tetrahedron has the maximum volume.

Gromov's norm is a seminorm on the real singular homology $H_3(\mathcal{M}; \mathbb{R})$ (here we are only considering the 3-homology). Let $[\mathcal{M}]$ denote the third homology class of the whole manifold (the fundamental class). Consider all singular 3-cycles c that represent this class, for example, the cycles coming from triangulations. We can express c as $c = \sum a_i\sigma_i$. Define $|c| = \sum |a_i|$. Then define the *Gromov norm* of $\mathcal{M}$

$$\|\mathcal{M}\| = \inf\{|c| : c \text{ is a singular cycle representing } [\mathcal{M}]\}.$$

(See [Thurston 1979, §6.1] or [Benedetti and Petronio 1992, §C.3].) The Gromov norm has the property that for any continuous map $f : \mathcal{M} \to \mathcal{M}_1$,

$$\|\mathcal{M}\| \geq |\deg f|\, \|\mathcal{M}_1\|.$$

Gromov's Theorem says that

$$\|\mathcal{M}\| = \frac{\mathrm{Vol}(\mathcal{M})}{V_3},$$

where $V_3 = 1.01294\cdots$ is the volume of the regular ideal tetrahedron.

This is an remarkable formula. One can view it as showing that the volume of $\mathcal{M}$ a topological invariant, or as giving a topological interpretation of the volume of closed hyperbolic manifolds. The same formula holds for hyperbolic n-manifolds, $n \geq 2$. (Prove it for $n = 2$, when the volume is $4\pi(g-1)$ and $V_2 = \pi$.)

For an orientable, closed manifold M^3, not necessarily hyperbolic, the number $V_3\, \|M^3\|$ is called the *simplicial volume* of M^3. The Gromov norm has the property that for any continuous map, say $f : \mathcal{M} \to \mathcal{M}_1$,

$$\|\mathcal{M}\| \geq |\deg f|\, \|\mathcal{M}_1\|.$$

Can the theory be extended to cusped manifolds of finite volume?

4-15. *The space of geodesics.* Show that the space of geodesics in $\mathbb{H}^2$, in the topology coming from the Hausdorff metric, is homeomorphic to the quotient $(\mathbb{S}^1 \times \mathbb{S}^1 \setminus \delta)/\langle J\rangle$, where J is the involution $(\zeta_1, \zeta_2) \to (\zeta_2, \zeta_1)$ and $\delta = \{(\zeta, \zeta),\ \zeta \in \mathbb{S}^1\}$ is the diagonal.

Show that the quotient space is in turn homeomorphic to an *open* Möbius band. Hint: Represent $\mathbb{S}^1 \times \mathbb{S}^1 \setminus \delta$ as a square torus less a simple loop representing the diagonal: take the torus to be the quotient of a square lattice, and in a fundamental square represent δ as a diagonal. Let Δ_1, Δ_2 denote the resulting triangles and J the reflection $J : \Delta_1 \leftrightarrow \Delta_2$ in δ. Note that Δ_1 is a fundamental domain for the torus

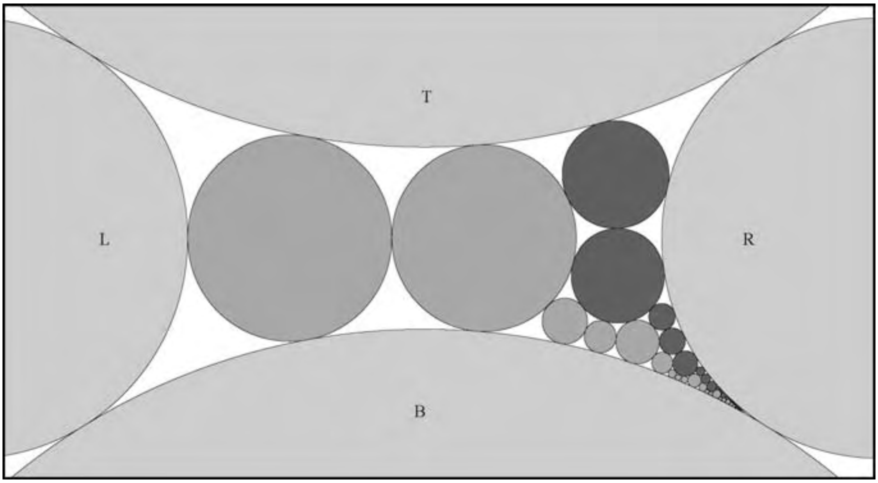


Fig. 4.3. Computation of the modulus of a marked quadrilateral.

group augmented by J. Label the sides s_1, s_2 of Δ_1 and s'_1, s'_2 of Δ_2 where $J : s_i \leftrightarrow s'_i$. Under the augmented group a pair of points on side s_1 is equivalent to a pair on s_2, in the opposite order. Using this correspondence, glue the side s_1 to s_2. This forms a Möbius band.

4-16. *Circle packings I.* In this problem we will report on Bob Brooks' important papers [1985; 1986]. We start by noting that three circles each externally tangent to the other two, bound a circular triangle — an ideal triangle (actually two of them, one containing ∞). In $\mathbb{S}^2$ such a configuration is uniquely determined up to Möbius equivalence.

Next consider four circles, each externally tangent to exactly two others. They bound a circular quadrilateral Q. Two such configurations are generally not Möbius equivalent. Recall that two rectangles of widths a, a' and heights b, b' are Möbius equivalent in such a way that the horizontal sides and vertical sides correspond if and only if $a/b = a'/b'$. Brooks discovered an analogous modulus for circular quadrilaterals:

Designate one pair of opposite sides of Q as *horizontal* and the other as *vertical*. Order the sides as top, bottom, right and left. This makes Q into a *marked quadrilateral*. A circle bounding a disk in Q will be called *horizontal* if it is tangent to the left vertical side and the two horizontal sides; it will be called *vertical* if it is tangent to the top horizontal side and to the right and left vertical sides.

Suppose, for example, we can insert a horizontal circle. Its exterior in Q consists of two circular triangles and perhaps another quadrilateral Q_1. Insert another horizontal circle in Q_1, if possible. Continue the process of inserting horizontal circles until after $n_1 \geq 1$ steps we can no longer do so (this process must stop after a finite number of times). Then with the remaining quadrilateral Q_{n_1}, start inserting vertical circles until that is no longer possible. Say the number of vertical circles is $n_2 \geq 1$. Then again start inserting horizontal ones. And so on. Form the continued fraction

$$r(Q) = n_1 + \cfrac{1}{n_2 + \cfrac{1}{n_3 + \cdots}};$$

it converges to a positive real number $r(Q)$ which is rational if and only if the sequence $\{n_i\}$ terminates after a finite number of terms.

Theorem 4.12.4. *Two marked quadrilaterals Q, Q' are Möbius equivalent consistent with the marking if and only if $r(Q) = r(Q')$.*

The set of marked Q with finitely many circles is dense in the configuration space of all marked quadrilaterals Q.

Next consider the case of a geometrically finite group G without parabolics. By a circle packing in Ω we will mean a G-invariant collection of round circles with mutually disjoint interiors which projects to $\partial\mathcal{M}(G) = \Omega(G)/G$ to give a finite packing by simple curves, which we will again call circles, bounding mutually disjoint regions. Add additional circles so as to arrive at a G-invariant packing in $\Omega(G)$, finite in $\partial\mathcal{M}(G)$ with the property that all *interstices* are either triangles or quadrilaterals. This will now be our definition of circle packings.

Each circle packing has a *nerve* which is the graph obtained by associating each circle with a vertex and each pair of tangent circles with an edge.

From the point of view of $\partial\mathcal{M}(G)$ there are a finite number of quadrilaterals. Mark the quadrilaterals and compute their moduli $\{r(Q)\}$.

Theorem 4.12.5. *Suppose G, G' are geometrically finite (torsion free) groups without parabolics. Assume there is an isomorphism θ between nerves of circle packings on $\partial\mathcal{M}(G)$ and $\partial\mathcal{M}(G')$ such that corresponding marked quadrilaterals have the same moduli. Then G and G' are conjugate; $\mathcal{M}(G)$ and $\mathcal{M}(G')$ are isometric.*

The group G may be quasiconformally deformed an arbitrarily small amount to a new group G_ϵ which has an isomorphic circle packing such that the moduli of the corresponding quadrilaterals on $\partial\mathcal{M}(G_\epsilon)$ are rational. Therefore G_ϵ has a finite circle packing such that all interstices are triangles.

There exists a geometrically finite group $\Gamma_\epsilon^{**} \supset G_\epsilon$ with $\mathcal{M}(\Gamma_\epsilon^{**})$ a closed manifold.

Here are some remarks concerning the proofs. Given a circle packing of $\Omega(G)$, the group Γ generated by G plus the reflections in the hyperbolic planes supported by the circles is geometrically finite. (It suffices to consider the orientation preserving index two subgroup.) The deformation space of Γ depends on one real parameter for each quadrilateral. Small changes in the parameter can be realized by a quasiconformal map with small dilatation that maps the circle packing onto a combinatorially identical one close by.

Once we get a circle packing such that the moduli of all quadrilaterals is rational, we can add circles to get a larger packing such that all interstices are triangles. Denote by Γ_ϵ the group generated by G_ϵ and reflection in all the planes supported by circles. Γ_ϵ too is geometrically finite.

In Ω_ϵ, put a new circle through the three vertices of each interstitial triangle. It is orthogonal to the three forming the triangle. Then form the even larger group Γ_ϵ^* generated by Γ_ϵ and reflections in the hyperbolic planes supported by the new circles. Again we may consider the orientation preserving subgroup, which is of index two. It too is geometrically finite. There are only a finite number of points which, mod Γ_ϵ^*, are not contained in the interior of any disk; these become rank two parabolic fixed points. The corresponding manifold has finite volume.

So far the proof can be adapted to apply when there are parabolics in G.

Now we outline the proof of the third statement of the theorem. We have the finite volume group Γ_ϵ^*, where we have replaced the original by its orientation preserving subgroup, now of index four in the group with all the circle reflections. The only parabolics are those associated with the rank two subgroups. Next do Dehn surgery on the rank two cusps to get a group Γ_ϵ^{**} whose quotient manifold is closed. Our original deformation G_ϵ is a subgroup of Γ_ϵ^{**} without any parabolics. Because Γ_ϵ^{**} can be taken arbitrarily close to Γ_ϵ^*, a small quasiconformal deformation G'_ϵ of G_ϵ appears as a subgroup of Γ_ϵ^{**}. This completes the argument.

4-17. *Circle packings II.* There is an important technique based on circle packings, motivated by results of Koebe, Andreev, and Thurston and by a conjecture of Thurston (confirmed in [Rodin and Sullivan 1987]), and extensively developed into a tool in pure and applied mathematics, especially by Ken Stephenson and colleagues [Stephenson 2005]. We will base our report here on [Beardon and Stephenson 1990]; see also the expositions [Stephenson 1999; 2003].

Suppose S is an abstract orientable, closed polyhedral surface, composed of triangles (actually one can begin with a topologically triangulated surface or even an abstract 2-complex). The prototypical example is a topological 2-sphere. There are likely to be vertices of both positive and negative discrete curvature. Denote the graph of edges and vertices by Γ. The basic theorem asserts that there exists a homeomorphic closed Riemann surface R — a sphere, a torus, or a hyperbolic surface — such that Γ is combinatorially isomorphic to the nerve of a circle packing of R. That is, there is a circle packing of R (in the spherical, euclidean or hyperbolic metric) such that the graph (nerve) formed by taking vertices to be the centers of the circles, and edges the geodesic segments between centers of circles, is isomorphic to Γ. Furthermore, R is uniquely determined by the combinatorics of the packing, up to conformal equivalence. For the case $g \geq 2$, the set of Riemann surfaces that can be circle packed is dense in the Teichmüller space [Bowers and Stephenson 1992].

If the original surface S is a region in $\mathbb{C}$, say with fractal boundary, one typically proceeds as follows. Restrict the regular hexagonal packing of $\mathbb{C}$ with radii δ to S, and select the connected component containing a prescribed basepoint O. Form the nerve of the resulting configuration in S and join the boundary vertices to ∞ by arcs, thereby creating a graph on $\mathbb{S}^2$. Construct the corresponding circle packing of $\mathbb{S}^2$. Matters can be normalized so that ∞ corresponds to ∞ and the disk in the packing with center ∞ is the exterior of the unit disk. In effect, the packing of S becomes a

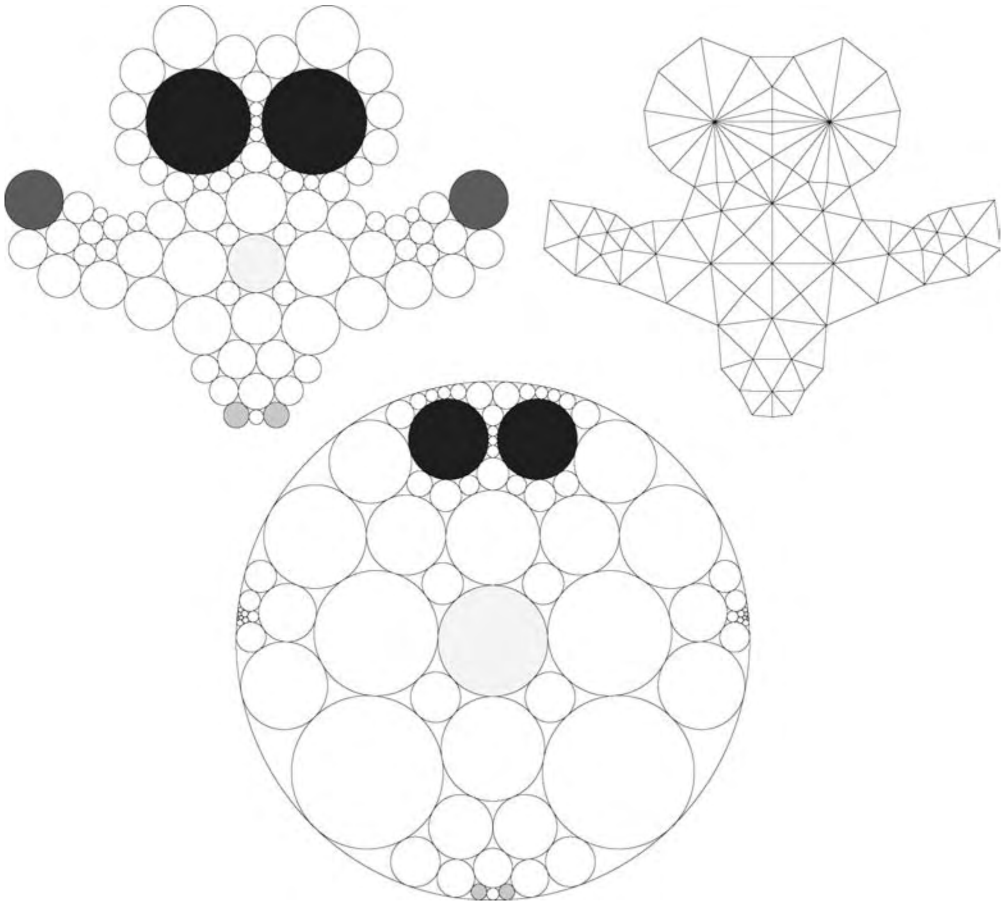


Fig. 4.4. Top: a triangulation of $\mathbb{S}^2$ (right) taken from an owl (left) by drawing edges between the centers of tangent circles and lines to ∞ from boundary circles. Bottom: its realization as a circle packing in the unit disk.

packing of the unit disk $\mathbb{D}$. The euclidean circles are equally hyperbolic circles, or horocycles, if they are tangent to $\partial\mathbb{D}$. It is proved in [Rodin and Sullivan 1987] that as the radius $\delta \to 0$, the quasiconformal simplicial map set up by mapping the nerve of the packing of S to that in $\mathbb{D}$ converges uniformly on compact subsets of S to a Riemann mapping (there needs to be another normalization to account for rotation about O). Many other investigations have followed from this pioneering result.

If S is a surface with boundary, there is a commonly used intrinsic method of assigning boundary values. Start a finite triangulation of S. One method is to assign the value one to each boundary edge, and require that the nerve of the circle packing have the same property. This gives a different shape to the circle packing of a plane region than the process above. Another method is to assign a "radius" to each boundary vertex and require the boundary circles in the resulting packing to have the assigned radii. This can be done in hyperbolic geometry where the packing is done in $\mathbb{D}$ and radius ∞ corresponds to a horocycle. In either case the circle packing is uniquely determined up to Möbius equivalence.

The circle packing method can be described as a process to change combinatorics into geometry! For example, experimental work in neurology uses this method in an attempt to set up a universal coordinate system for the cortex surface of the human brain. The surface of the cerebral cortex is highly convoluted and varies from person to person. Using MRI scans it can be triangulated into many tiny triangles. The circle packing method can be used to replace the triangulated cortex by a circle packing of a plane region. The hope is that this "flattening" of the cortex will make those of different individuals more easily comparable.

There is a theory of packings where the circles have prescribed intersection angles given by a function on the edges of the triangulation [Bobenko and Springborn 2004; Rivin 1996; Thurston 1979, Chapter 13]. ,It is closely tied up with the study of hyperbolic polyhedra in $\mathbb{H}^3$.

4-18. *Ideal triangulations; spinning*. Start with a closed hyperbolic surface S of genus g. Take a family P of $3g-3$ mutually disjoint simple geodesics forming a *pants* decomposition of $2g-2$ triply connected regions (see Section 5.3). On each geodesic of P fix one point. Draw 3 mutually disjoint geodesic arcs between the pairs of points, dividing each pants into two simply connected regions. Now "spin" the "triangulation" by applying Dehn twists (Exercise 5-6) of higher and higher order about the simple loops of P. At each stage realize the edges of the auxiliary arcs by geodesic segments. In the limit there will result a geodesic lamination Λ of S. The leaves of Λ will consist of the geodesics of P, plus a finite number of infinite length leaves spiraling about these, three in each P. Up in the universal cover, say $\mathbb{D}$, the lifts will give a tessellation by ideal triangles. Each ideal triangle will project to a complementary component of Λ; a finite number will cover S, except for a set of measure zero. In fact there will be exactly $2(2g-2)$ triangles.

Return again to S with the geodesics giving a pants decomposition P. Suppose $\theta : \pi_1(S) \to \pi_1(\mathcal{M})$ is an injection to some hyperbolic manifold, parabolics corresponding to parabolics. Each geodesic of P corresponds to a uniquely determined geodesic of $\mathcal{M}$. Even more, each geodesic of Λ corresponds to a unique geodesic of $\mathcal{M}$. Now fill in the spaces with (immersed) ideal triangles. This results in a pleated surface in $\mathcal{M}$. It is easiest to carry out this construction in $\mathbb{H}^3$ directly — there the ideal triangles are embedded. In visualizing the result in $\mathcal{M}$, recall the uniform injectivity property (page 150).

It is much easier to find ideal triangulations when there are punctures. Designate a point v on S. We can divide S into $4g-2$ triangles were all the vertices are at v. The triangulation has $6g-3$ edges. Now puncture S at v. In the hyperbolic metric on the punctured surface, in the same combinatorics we can take all the edges to be geodesics, dividing the surface into $4g-3$ ideal triangles. (Correspondingly, the universal cover will be tessellated by ideal triangles.) We can talk about the space $\mathcal{C}(S)$ of all possible ideal triangulations based on v.

Once an ideal triangulation is specified on S it is specified at all points of $\mathrm{Teich}(S)$.

Return now to the ideal triangulation on the once punctured surface S. It is determined by a real vector $\vec{x} \in \mathbb{R}_+^{6g-3}$ as follows. To each edge e_k associate a positive number: Focus on the two ideal triangles that share e_k and take the orthogonal projection of the third ideal vertex of each onto e_k. Take the k-th component of $\vec{x}$ to be e^{x_k} where $x_k \geq 0$ is the distance between the two projections. In fact, once we fix a basepoint, there is a unique left earthquake (Exercise 3-31) that realizes the described motion. The vectors $\{\vec{x}\}$ give a coordinate system for $\mathcal{C}(S)$ over $\mathrm{Teich}(S)$. It is no accident that $(6g-3)$ is the real dimension of $\mathrm{Teich}(R)$ up to normalization. It is shown in [Bonahon 1996] how to get real analytic coordinates of $\mathrm{Teich}(R)$ by using these numbers.

According to Jeff Weeks (personal communication) it is "almost surely true" that every finite volume hyperbolic manifold with cusps can be decomposed into positively oriented ideal tetrahedra, but at this writing this remains a conjecture. However it is true that every such manifold has a decomposition into ideal polyhedra (Exercise 3-25). Even so, there is no proof that one can subdivide the polyhedra into ideal tetrahedra so that the subdivisions agree on the common faces of adjacent polyhedra [Weeks 1993; 2005; Petronio and Weeks 2000].

On the other hand, the decomposition can be achieved if the "positively oriented" requirement is dropped to allow some flattened tetrahedra; the flattened tetrahedra appear where the polyhedral subdivisions do not agree on a common face. Here a flattened tetrahedron is a planar quadrilateral whose four ideal vertices lie on a circle in $\mathbb{S}^2$. SnapPea [Weeks n.d.] uses such decompositions to numerically approximate the volumes of cusped manifolds, using the formulas recorded in Exercise 1-23. More generally, the program computes hyperbolic structures on manifolds, if such exist. It is a fundamental and productive tool in experimental work on the subject. The process SnapPea searches for a hyperbolic structure is as follows [Weeks 2005]. Start with a topological ideal triangulation τ of n simplices of a cusped manifold, for example, a knot complement. Using the combinatorial data implicit in τ, write down a system of equations expressing the fact that a solution allows τ to be homotoped to a (geodesic) ideal triangulation that is compatible with the given hyperbolic structure.

In more detail, a (geodesic) ideal tetrahedron is uniquely determined up to isometry by the cross ratio of its ideal vertices (Exercise 1-22). To form the manifold, the ideal tetrahedra need to be put together by a sequence of face identifications. The total angle about each edge of the complex needs to be 2π. At the cusps there must be horosphere cross sections so as to become rank two cusps (in more general cases, there is a condition that in effect says that a Dehn surgery must be undertaken). The bottom line is that there results a finite number of algebraic conditions to be satisfied by the tetrahedra if the required identification can be accomplished (in $\mathbb{H}^3$). If SnapPea finds a solution, it is shown to be mathematically correct using Snap [Goodman n.d.], another software tool that uses exact arithmetic. See [Neumann 1999; Coulson et al.

2000; Neumann and Reid 1992] for more details. As indicated above, there may be some complications in the construction. See also Exercise 3-5.

There is a more complicated procedure used for closed hyperbolic manifolds.

4-19. Prove that $\mathrm{Int}(\mathcal{M}(G))$ has finite volume if and only if the injectivity radius $\mathrm{Inj}(x) \to 0$ as any $x \in \mathrm{Int}(\mathcal{M}(G))$ approaches the ideal boundary.

4-20. *Simple loops in $\mathcal{M}(G)$; primitive curves.* What should it mean that $\gamma \subset \mathcal{M}(G)$ is a simple loop, assuming γ is not retractable to a point or a cusp? On one level the answer is obvious: Every closed curve can be deformed in space so it becomes a simple loop. But on a deeper level we may want to use the criterion that arises as follows: Given a closed curve γ let $g \in G$ be an element determined by γ; that is, the lift of γ from a point $O \in \mathbb{H}^3$ over its origin terminates at $g(O)$. Here g is necessarily loxodromic. The stabilizers of the pair of fixed points of g determine a cyclic subgroup. Is g a generator of this cyclic group? Or, is there another element g_1 of it such that $g = g_1^n$ with $n \geq 2$? If g is indeed a generator, the curve γ is called *primitive*. Even if γ is freely homotopic to a simple loop on $\partial\mathcal{M}(G)$, it may not be primitive. Of course, for a surface itself, all nontrivial simple loops are primitive.

Here is an example. In Exercise 4-24 we will show that there is a $\mathcal{M}(G)$ with a single rank two cusp and an essential cylinder C with one boundary component on a cusp torus and the other in $\partial\mathcal{M}(G)$. Let $c \subset C$ be a central curve. Let c' denote the freely homotopic curve on the solid torus. Let d be a simple loop on the cusp torus with is transverse to c'. Do $(1, 2)$ Dehn surgery on the cusp so that in the new manifold which has no rank two cusps, $c' \sim d^2$. In the new manifold, c remains freely homotopic to a simple loop on the boundary, but it is not primitive.

In the original manifold there is also an essential cylinder C^* with both boundary components on $\partial\mathcal{M}(G)$ and whose central curve is freely homotopic to c. In the new manifold, C^* remains an essential cylinder, and its core curve is not primitive.

On the other hand, a simple closed *geodesic* in $\mathcal{M}(G)$ is automatically primitive!

4-21. *Geometric limits by renormalization.* Suppose $\{\gamma_n\}$ is a collection of closed geodesics with the property that only a finite number meet any given compact set $K \in \mathcal{M}(G)$ and that their lengths are uniformly bounded by $L < \infty$. Fix origins $O_n \in \gamma_n$. Consider the sequence $\{T_n G T_n^{-1}\}$ where T_n maps O_n to the origin O in say the ball model. Prove that there is a geometrically convergent subsequence to a limit H. Show that in $\mathcal{M}(H)$, the renormalized geodesics $\{\gamma'_m\}$ converge to a closed geodesic γ.

4-22. *Every surface has a decomposition by pants of medium size.* Consider the Teichmüller space Teich_g of closed surfaces of genus g. Prove with Bers [1985] that there exists a number L_g such that every surface $S \in \mathrm{Teich}_g$ has a pants decomposition (see Exercise 4-18 and Section 5.3) in which no boundary curve of a pants has length exceeding L_g.

4-23. If Λ is not uniformly perfect (Exercise 1-30), prove that there is a sequence of Möbius transformations $\{A_n\}$ such that $\{A_n\}$ converges in the Hausdorff topology to a closed set that contains an isolated point. Conversely, if there is such a sequence $\{A_n\}$, Λ is not uniformly perfect.

4-24. *Joining unpaired or paired punctures.* Here are two related constructions.

Construction 1. Suppose h_1, h_2 are horodisks at different parabolic fixed points corresponding to two rank one parabolic subgroups of a group G (or of two different groups G_1, G_2).

Take any Möbius transformation T that sends the exterior of h_2 onto the interior of h_1 and conjugates the parabolic subgroups. Form the augmented group $G^* = \langle G, T \rangle$.

Down in the quotient, there is a horocycle bounding a once punctured disk about each of the punctures. Join the manifold to itself by pasting together these two punctured disks. The punctured disk appears in the interior of the new manifold $\mathcal{M}(G^*)$, which is geometrically finite if the original manifold is. A single rank one parabolic conjugacy class arises from the conjugation of the two possibly different original classes. See Figure 4.5.

Construction 2. Suppose h_1, h_2 are disjoint horodisks at a rank one parabolic fixed point.

Take any Möbius transformation T that sends the exterior of h_2 onto the interior of h_1 and conjugates the rank one parabolic group to itself. Form the augmented group $G^* = \langle G, T \rangle$.

Down in the quotient, the horocycles about two punctured disks form the boundary of a cusp pairing cylinder C. The effect of adjoining T is to paste together the two punctured disks and the horocycles bounding them. This creates a solid cusp torus in $\mathcal{M}(G^*)$ which corresponds to the rank two group that has been created. The new manifold is geometrically finite if the initial one is.

The simplest application of Construction 2 is to the modular group M_2 of Exercise 2-9, which acts in the upper and lower half-planes. The new parabolic has the form $Tz = z + \tau$ where $\operatorname{Im} \tau \neq 0$. The process gives us a 4-punctured sphere with

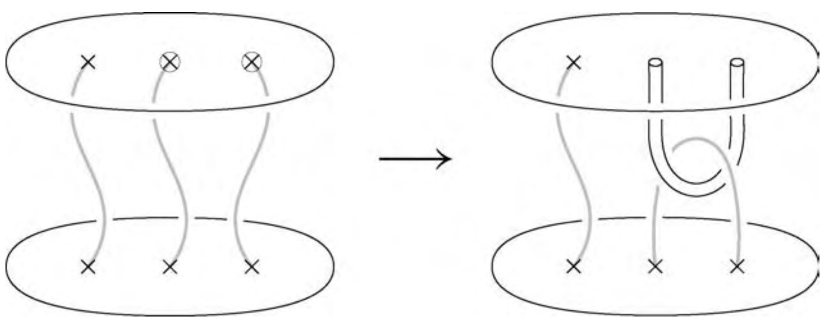


Fig. 4.5. Two unpaired punctures determine a handle. The gray curves represent solid pairing tubes.

complex parameter τ. For this case we can take τ so that $|\mathrm{Im}\,\tau| \geq 1+2\epsilon$. The resulting group has infinitely many regions of discontinuity, each of which is simply connected. The manifold $\mathcal{M}(G^*)$ is bounded by a 4-punctured sphere with the punctures arranged in two pairs. In addition $\mathcal{M}(G^*)$ has a rank two cusp which is hidden from the viewer looking only at $\mathbb{S}^2$. A simple loop that separates the two punctures of each pair on $\partial\mathcal{M}(G^*)$ is freely homotopic to a simple loop on the cusp torus.

Construction 2 can be used to give an explicit example of the following situation referred to in Exercise 4-20: Suppose $\mathcal{M}(G)$ has a rank two cusp and let $\mathcal{M}_0$ be the result of removing a corresponding solid cusp torus from $\mathcal{M}(G)$. It is possible that there are two nontrivial, nonparallel simple loops α_1, α_2 in $\partial\mathcal{M}(G)$ and a simple loop α_c in the cusp torus of $\mathcal{M}_0$, such that the pairs (α_c, α_1), and (α_c, α_2) both bound essential cylinders in $\mathcal{M}_0$. Of course, (α_1, α_2) then bound an essential cylinder in $\mathcal{M}(G)$.

To obtain such an example, start with the algebraic limit of a quasifuchsian group, say closed surface group, that has a surface of genus g on the bottom, and a pinched surface — say a single pinched surface — on top. There is a simple loop α_2 in the bottom surface that bounds an essential cylinder with a loop about either of the punctures on top. Now apply Construction 2 to replace the top pinched surface by a closed surface S of genus g and a rank two cusp. The simple loops about the punctures determine a simple loop α_1 on S, which is freely homotopic to α_2. The two simple loops about the punctures also determine the meridian α_c on a cusp torus corresponding to the new rank two cusp. These three simple loops fulfill the requirements.

Figure 4.6 is an implementation of Construction 2. The starting point is the thrice punctured sphere fuchsian group. The 6 punctures are arranged in 3 pairs, each pair supports two horodisks. The first picture shows the result of joining one pair of horodisks resulting in the quotient manifold $\mathcal{M}_1$ with one rank two cusp and $\partial\mathcal{M}_1$ a 4-punctured sphere. For the second picture, another pair of horodisks are joined

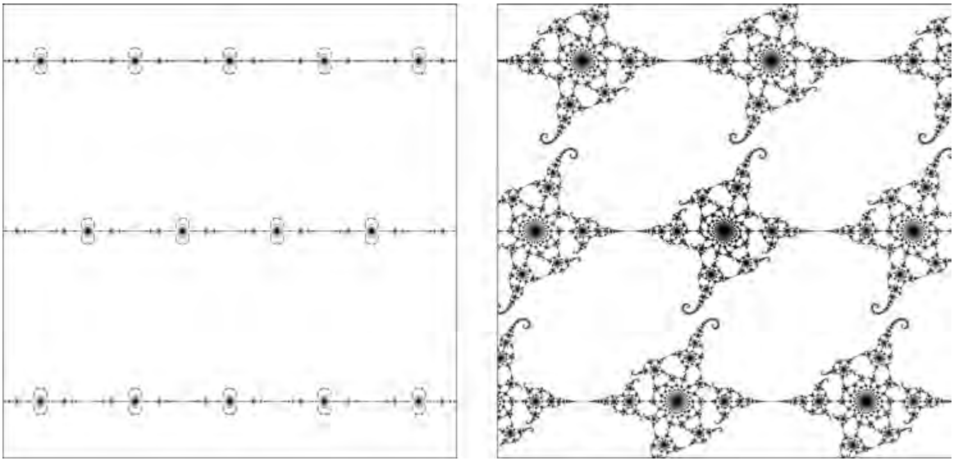


Fig. 4.6. Earle–Marden coordinates for 4-punctured spheres and 2-punctured tori.

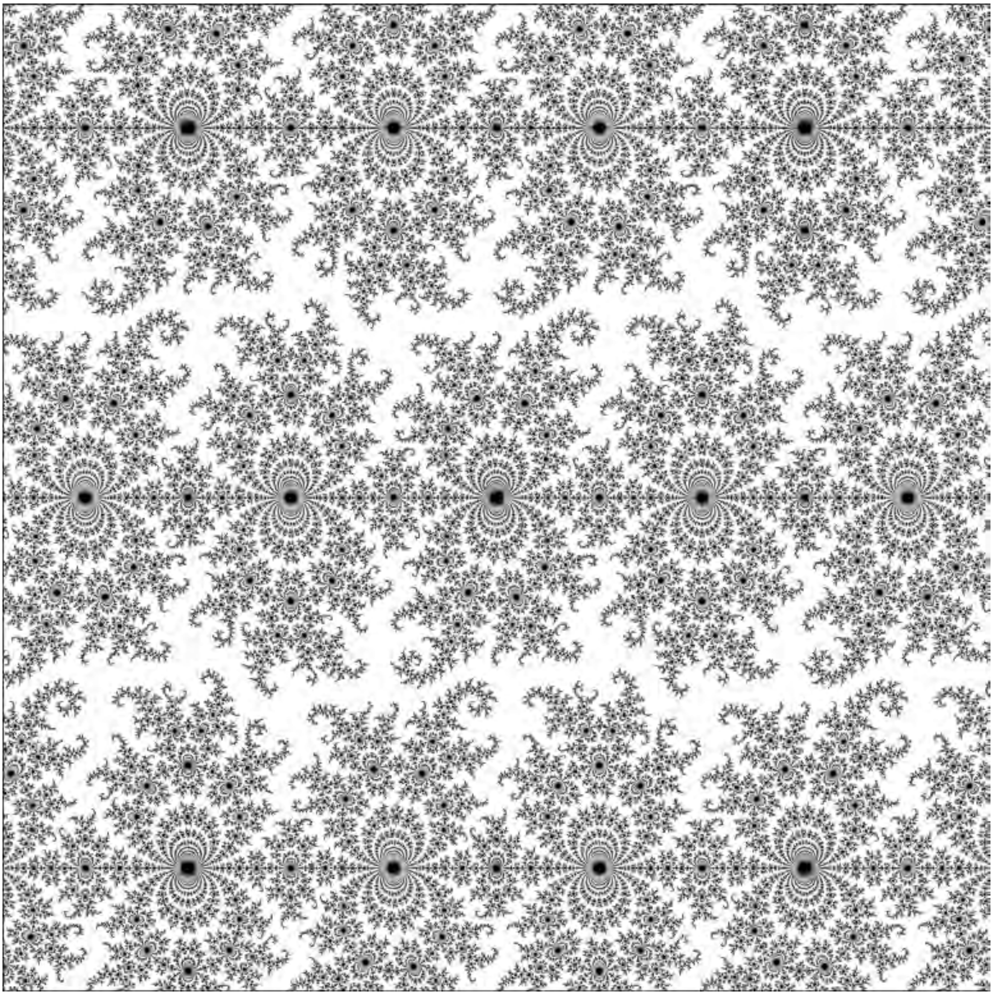


Fig. 4.7. Earle–Marden coordinates for genus-2 surfaces.

resulting in $\mathcal{M}_2$ with two rank two cusps and $\partial\mathcal{M}_2$ a 2-punctured torus. The final step gives Figure 4.7. Here the final pair of horodisks are joined resulting in $\mathcal{M}_3$ with three rank two cusps and $\partial\mathcal{M}_3$ a genus-two closed surface. There is one free complex parameter for each rank-2 cusp. These extend to become holomorphic coordinates of the Teichmüller spaces of the respective surfaces and are called the *Earle–Marden coordinates*. In this sequence of figures, the parameters are chosen so that each group is a subgroup of the next.

5

Deformation spaces and the ends of manifolds

Our work in the earlier chapters, especially our understanding of the structure of geometrically finite manifolds, has prepared the ground for understanding the results that will be discussed here, without most proofs. At center stage are the three fundamental conjectures, now theorems, concerning the structure of hyperbolic manifolds with finitely generated fundamental groups but which are not geometrically finite: the Tameness Theorem, Ending Lamination Theorem and Density Theorem. The chapter begins with a study of the representation variety. We go on to present the quasiconformal deformation spaces of geometrically finite groups and their boundaries, in particular the quasifuchsian spaces and their Bers slices. To understand generic boundary points of these spaces we are confronted with the need to understand geometrically infinite ends. These are the ends with "missing" Riemann surface boundary components. However tameness tells us that these ends have product neighborhoods. And instead of boundary components there are "ending laminations". Then density tells us that indeed every finitely generated kleinian group is the algebraic limit of geometrically finite groups.

5.1 The representation variety

Suppose $G=\langle g_1, g_2, \dots, g_r\rangle$ is a finitely generated kleinian group (without elliptics), that is G is isomorphic to the fundamental group of $\mathcal{M}(G)$. By the Scott–Shalen theorem, there are a finite number of relations $\{R_k(g_1, \dots, g_r)=\mathrm{id}\}$, each of which is a word in the generators, such that any relation in the group is a consequence of these.

If we vary the entries in the generating matrices $\{g_i\}$ we will get a new set of Möbius transformations. The group generated by the new set will be a homomorphic image of G if and only if the new set of generators $\{g_i'\}$ satisfy the same relations $\{R_k(g_1', \dots, g_r')=\mathrm{id}\}$. These are algebraic equations in the matrix entries. In this book, a *representation* of G is a homomorphism into $\mathrm{PSL}(2,\mathbb{C})$ or into $\mathrm{PSL}(2,\mathbb{R})$.

A representation is called *elementary* if the image group is elementary and called *reducible* if the image group has a common fixed point (Exercise 2-1).

We will not want to distinguish between conjugate representations. Therefore, for our purposes, we define the *representation variety* $\mathfrak{R}(G)$ as the quotient of the set of *nonelementary* representations under the conjugation equivalence relation $\equiv$: two representations satisfy $\varphi_1 \equiv \varphi_2$ if there is a Möbius transformation T such that $\varphi_2(g) = T \circ \varphi_1(g) \circ T^{-1}$, for all $g \in G$. Thus

$$\mathfrak{R}(G) = \{\varphi \mid \varphi : G \to H \text{ is a type preserving homomorphism to a nonelementary group}\}/\equiv.$$

By *type preserving* we will mean that if $g \in G$ is parabolic, so is $\varphi(g)$. This means that rank one and rank two parabolic groups are preserved, but it does not prevent new parabolics (or elliptics) appearing in the target groups.

We have to admit that we are taking certain liberties in our use of the term "representation variety". Actually $\mathfrak{R}$ is the quotient of an open subset of the affine algebraic variety of type preserving representations into $\mathrm{SL}(2, \mathbb{C})$. For a discussion of representation varieties see [Culler and Shalen 1983] or [Kapovich 2001].

We emphasize that $\mathfrak{R}(G)$ is not just a space of groups, but a space of *marked groups*. A group H which the target of nonconjugate homomorphisms φ_1, φ_2, $G \to H$, is represented by *distinct* points of $\mathfrak{R}(G)$.

Each (normalized) matrix depends on 3 complex parameters; the set of generators depends on $3r$ complex parameters. That s relations must be satisfied costs $3s$ conditions giving a dimension of at most $3r - 3s$. But then we do not distinguish between two groups that are conjugate. Conjugacy depends again on 3 complex parameters, so the dimension of the quotient is now at most $3r - 3s - 3$.

The parabolicity condition further reduces the dimension of $\mathfrak{R}(G)$. Each cyclic parabolic conjugacy class gives rise to a relation of the form $\mathrm{tr}^2(R'_k(g_1, \dots, g_r)) = 4$. If there are b_1 rank one classes, and b_2 rank two classes, $\mathfrak{R}(G)$ has complex dimension $3r - 3s - b_1 - 2b_2 - 3$. So if G is a fuchsian group representing a closed surface of genus g, the $\mathbb{C}$-dimension of $\mathfrak{R}(G)$ is $3 \cdot 2g - 3 - 3 = 6g - 6$. If G represents a finite area surface of genus g and b punctures the dimension count is $3(2g+b-1)-b-3 = 6g+2b-6$. Also see Exercise 4-13. For rigorous computations see [Culler and Shalen 1983], or the discussion of $\mathfrak{T}(G)$ below.

The groups represented by points of $\mathfrak{R}(G)$ are in general not discrete and perhaps not even finitely presented. Furthermore, $\mathfrak{R}(G)$ is typically not connected. For example if X is an orientation reversing Möbius transformation and there is no orientation preserving map that induces a representation from G to XGX^{-1}, then G and XGX^{-1} are in different components.

We have removed the elementary groups. Of these, the most significant are subgroups conjugate into $\mathrm{SO}(3)$, rotations of the ball. For example there is an isomorphism of a closed surface group of any genus ≥ 2 to a subgroup of $\mathrm{SO}(3)$ [Greenberg 1981]. Besides fitting better into our theory, a big advantage of removing them is that the quotient is then Hausdorff: disjoint points have disjoint neighborhoods. Not only for nonelementary groups, but more generally for irreducible groups, the Hausdorff property holds, as shown by the following Lemma. Denote the space of irreducible representations by $\mathrm{Hom}_{\mathrm{ir}}(G; \mathrm{PSL}(2, \mathbb{C}))$.

In all these spaces, the topology used is the topology of algebraic convergence.

Lemma 5.1.1. *Assume that for* $\varphi_1, \varphi_2 \in \mathrm{Hom}_{\mathrm{ir}}(G; \mathrm{PSL}(2, \mathbb{C}))$ *the target groups* $H_1 = \varphi_1(G)$ *and* $H_2 = \varphi_2(G)$ *are not conjugate. There exist neighborhoods* N_1 *of* H_1 *and* N_2 *of* H_2 *in* $\mathrm{Hom}_{\mathrm{ir}}(G; \mathrm{PSL}(2, \mathbb{C}))$ *such that no conjugate of* $\phi \in N_1$ *is in* N_2.

Proof. Given a neighborhood $N_1 \subset \mathrm{Hom}_{\mathrm{ir}}(G; \mathrm{PSL}(2, \mathbb{C}))$ of H_1 with $H_2 \notin N_1$, we claim that there exists a neighborhood N_2 of H_2 with the property that no representation $\phi \in N_1$ is conjugate to a representation in N_2. If we can prove this, by reversing the roles of H_1 and H_2, both neighborhoods can be taken sufficiently small to meet the requirements.

Assume our claim is false. Then there is a sequence $\phi_n \in \mathrm{Hom}_{\mathrm{ir}}(G; \mathrm{PSL}(2, \mathbb{C})$ in N_1 and a corresponding sequence of Möbius transformations $\{T_n\}$ such that the representations

$$\phi'_n(G) = T_n \circ \phi_n(G) \circ T_n^{-1},$$

satisfy $\lim \phi'_n(G) = \varphi_2(G) = H_2$.

By Lemma 2.1.1 we may assume the fixed point(s) (p_n, q_n) of T_n converge to $(p, q) \in \mathbb{S}^2$, and that for all indices T_n is loxodromic with q_n the attracting fixed point, parabolic $p_n = q_n$, or elliptic. By hypothesis, no subsequence of $\{T_n\}$ can converge to a proper Möbius transformation.

By Exercise 2-1 there exist $h_1 = \varphi_1(g_1), h_2 = \varphi_1(g_2) \in H_1$ with distinct fixed points. We may assume that it is h_1 that does not fix p the limit of the repelling fixed points of $\{T_n\}$, assuming that the T_n are loxodromic — the other cases are handled similarly.

Choose a small enough neighborhood N_p of p so that $\phi_n(g_1)(N_p) \cap N_p = \varnothing$ for all large indices. Choose $z \notin N_p$. Then choose a small neighborhood N_q of q. If $p = q$ take $N_q \subset N_p$.

For all large indices, $T_n^{-1}(z) \in N_p$. So $\phi_n(g_1)(T_n^{-1}(z)) \notin N_p$. In fact for all large indices $T_n \phi_n(g_1) T_n^{-1}(z) \in N_q$. Since N_q can be taken arbitrarily small, we conclude that $\lim T_n \phi_n(g_1) T_n^{-1}(z) = q$, and then that this holds for all $z \neq p$. We have reached a contradiction to our assumption that $\lim \phi'_n(G) = H_2$. □

For more on representation varieties see Exercises 5-19 and 5-20.

The discreteness locus

The *discreteness locus* of a geometrically finite G is the following *closed* subset:

$$\mathfrak{R}_{\mathrm{disc}}(G) = \{\theta \in \mathfrak{R}(G) \mid \theta : G \to H \text{ is a (type preserving) isomorphism to a discrete group } H\}.$$

The target group $\theta(G)$ may be a quasiconformal deformation of G, but the isomorphism θ itself might *not* be induced by a quasiconformal map of $\mathbb{S}^2$. Thus the interior of the discreteness locus may have many components.

In the literature the commonly used notation for $\mathfrak{R}_{\text{disc}}(G)$ is $\mathcal{AH}(G)$, where $\mathcal{H}$ stands for homotopy equivalence and $\mathcal{A}$ reminds us that the topology is that of algebraic convergence.

Curiously, the convex cores of groups in $\mathfrak{R}_{\text{disc}}(G)$ are of a uniform "size", as indicated below.

Theorem 5.1.2.

(i) [Canary 1996] *The injectivity radius about points in the convex core of any given manifold $\mathcal{M}(H)$ is bounded above.*

(ii) [Evans 2006] *There exists $K = K(G)$ such that the radii of balls embedded in the convex core $\mathcal{C}(\mathcal{M}(H))$ of every $H \in \mathfrak{R}_{\text{disc}}(G)$ do not exceed $K(G)$.*

(iii) [Evans $\geq$ 2007] *Suppose $\mathcal{M}(G)$ has incompressible boundary. There exists $K = K(G) < \infty$ such that the injectivity radius of points in the convex core of $\mathcal{M}(H)$ is bounded by K for every $H \in \mathfrak{R}_{\text{disc}}(G)$.*

In the statement of (i), we are incorporating the Tameness Theorem. Note that the largest embedded balls about points of the convex core are not necessarily contained in the core. That the results are true was expected since there are pleated surfaces of fixed finite genus that exit each end. That the embedded balls are uniformly bounded prevents the manifolds from flying apart at the ends. The recently announced (ii) was a conjecture of McMullen. When $\mathcal{M}(H)$ is totally degenerated, the convex core coincides with $\mathcal{M}(H)$ itself. Statement (iii) does not hold for manifolds with compressible boundary. A counterexample is constructed in [Evans 2006] using a connected sum of surfaces.

On the other hand another conjecture of McMullen is still open: The rank of a group is the minimal number of generators. Given k, does there exist $R = R(k)$ such that for any closed manifold $\mathcal{M}(G)$ of rank k, every point $x \in \operatorname{Int} \mathcal{M}(G)$ has injectivity radius $\leq R(k)$? Note that if this were to fail, there would be a sequence whose geometric limit is $\mathbb{H}^3$; see [Souto 2006].

The quasiconformal deformation space

The group H is a *quasiconformal deformation* of the geometrically finite (nonelementary) group G if there is a quasiconformal map $F : \mathbb{S}^2 \to \mathbb{S}^2$ that induces an isomorphism $\theta : G \to H$ for which $F \circ g(z) = \theta(g) \circ F(z)$, for all $g \in G,\ z \in \mathbb{S}^2$. Such a group H is necessarily discrete and nonelementary. In particular, $g \in G$ is parabolic if and only if $\theta(g) \in H$ is parabolic. The map F in turn is uniquely determined up to normalization by its Beltrami differential defined on $\Omega(G)$ with the invariance property given by Equation (2.5). When we explicitly *normalize*, we will normally arrange things so that $(0, 1, \infty) \subset \Lambda(G)$ and F fixes these points.

The *quasiconformal deformation space* of G is defined as the following open, connected subset of the interior of $\mathfrak{R}_{\text{disc}}(G)$:

$$\mathfrak{T}(G) = \{\theta \in \mathfrak{R}(G) \mid \varphi \text{ is induced by a quasiconformal deformation of } G\}.$$

Thus two normalized deformations $F_1 \sim F_2$ are taken to be equivalent if they induce the same isomorphism θ. Another way of putting this is that if f_1, f_2 denote their projections to $\partial\mathcal{M}(G) \to \partial\mathcal{M}(\varphi(G))$, then $f_2 \circ f_1^{-1}$ extends to $\mathcal{M}(G) \to \mathcal{M}(G)$ and is homotopic to the identity on $\operatorname{Int}\mathcal{M}(G)$ — see Theorem 3.7.4. It is not necessarily true that $f_2 \circ f_1^{-1}$ is then homotopic to the identity on $\partial\mathcal{M}(G)$; see Exercise 5-25.

For a geometrically finite group G, one can follow its deformations by changes in a fundamental polyhedron $\mathcal{P}_{\mathcal{O}}(G)$ [Marden 1974a]. Along with any small deformation of G (or of any other point of $\mathfrak{T}(G)$) $\mathcal{P}_{\mathcal{O}}(G)$ is correspondingly deformed. When parabolics are preserved, no essential change occurs at the cusps. It follows that parabolic preserving homomorphisms of G close to the identity are actually isomorphisms induced by quasiconformal deformations of maximal dilatation close to one (*strong stability*: see [Marden 1974a]). Therefore $\mathfrak{T}(G)$ is an open subset of $\operatorname{Int}\mathfrak{R}_{\text{disc}}(G)$. It is a connected subset containing the identity because if μ is a Beltrami differential so is $t\mu$ for any $t \in \mathbb{C}$, $|t| < 1$. (If G is geometrically infinite, $\mathfrak{T}(G)$ is not open as nearby groups are nondiscrete.) When G is fuchsian, $\mathfrak{T}(G) = \operatorname{Int}\mathfrak{R}_{\text{disc}}(G)$.

If all components of $\Omega(G)$ are simply connected, that is, if $\partial\mathcal{M}(G)$ is incompressible, then $\mathfrak{T}(G)$ is a complex manifold biholomorphically equivalent to the product of the Teichmüller spaces of the individual components of $\partial\mathcal{M}(G)$. For then each component serves as the universal cover of the associated quotient.

On the other hand, if the components $\{\Omega_i\}$ over some S_i are not simply connected, there is a little problem because a conformal change of S_i may not cause a deformation in G. A simple loop γ around a handle of S_i may be compressible so a deformation in itself of the element of $\pi_1(S_i)$ corresponding to γ may result in no change to G. For an example, apply a Dehn twist (Exercise 5-6) to a compressing loop on the boundary of a handlebody.

The way out of this conundrum is given by Theorem 5.1.3. It was originally proved at the level of Beltrami differentials using "strong stability" [Marden 1974a] by Bers [1970b], Maskit [1971], and Kra [1972], with a three-dimensional interpretation in [Marden $\geq$ 2007]. Details of the proof are in Exercise 5-25.

Theorem 5.1.3. *Suppose G is geometrically finite. Denote the components of* $\partial\mathcal{M}(G)$ *by* $\{S_i\}$. *Then*

$$\mathfrak{T}(G) = \operatorname{Teich}(S_1)/\operatorname{Mod}_0(S_1) \times \cdots \times \operatorname{Teich}(S_k)/\operatorname{Mod}_0(S_k).$$

Here $\operatorname{Mod}_0(S_i)$ *is the fixed point free subgroup of biholomorphic automorphisms of* $\operatorname{Teich}(S_i)$ *generated by automorphisms of* S_i *that have an extension homotopic to the identity in the interior of* $\mathcal{M}(G)$.

The product $\operatorname{Teich}(S_1) \times \cdots \times \operatorname{Teich}(S_k)$ *is the universal cover of* $\mathfrak{T}(G)$; *the product equals* $\mathfrak{T}(G)$ *if and only if* $\partial\mathcal{M}(G)$ *is incompressible. The spaces* $\mathfrak{T}(G)$ *and* $\operatorname{Teich}(S_i)/\operatorname{Mod}_0(S_i)$ *are complex analytic manifolds*; $\mathfrak{T}(G)$ *has dimension*

$$\sum_{i=1}^{m}(3g_i + n_i - 3),$$

where g_i is the genus of the i-th component of $\partial\mathcal{M}(G)$ and n_i is the number of its punctures.

If $\mathcal{M}(G)$ is geometrically infinite and boundary incompressible, Theorem 5.1.3 still holds; in this case $\mathrm{Mod}_0(R_i) = \mathrm{id}$. For example, if $\Omega(G)$ is connected and simply connected (a singly degenerate group in $\partial\mathfrak{B}(G)$, see Section 6.1), then $\mathfrak{T}(G)$ has dimension $(3g + n - 3)$. In contrast if G_1 is fuchsian with $\mathbb{H}^2/G_1$ conformal to $\Omega(G)/G$ then $\mathfrak{T}(G_1)$ has dimension $(6g + 2n - 6)$. If all components R_i are triply punctured spheres, G is quasiconformally rigid, geometrically finite or not.

5.2 Homotopy equivalence

Understanding of isomorphisms to discrete groups in $\mathfrak{R}(G)$ involves the notion of *homotopy equivalence*. A homotopy equivalence between two manifolds M_1, M_2 is a pair of continuous mappings $f_1 : M_1 \to M_2$ and $f_2 : M_2 \to M_1$ such that $f_2 \circ f_1 : M_1 \to M_1$ is homotopic to the identity and $f_1 \circ f_2 : M_2 \to M_2$ is homotopic to the identity. In particular M_1 and M_2 have isomorphic fundamental groups.* For example, the 3-manifold $S \times (0, 1)$ is homotopy equivalent to the surface S; the one-holed torus is homotopy equivalent to the three-holed sphere. A hyperbolic manifold is homotopy equivalent to its compact or relative compact core. For a general development see [Johannson 1979].

Conversely, it is well known in topology [Whitehead 1978, Theorems 3.5, 7.1] that two manifolds whose higher homotopy groups vanish (as is the case for hyperbolic manifolds) are homotopy equivalent if (and only if) they have isomorphic fundamental groups.

Theorem 5.2.1 [Swarup 1980]. *Suppose $\mathcal{M}(G)$ is geometrically finite and $\mathcal{M}_0(G)$ the compactification resulting from removing solid pairing tubes and solid cusp cylinders. If $\partial\mathcal{M}_0(G)$ is not empty, there are only a finite number of compact manifolds $\mathcal{M}_0(H)$ that are homotopy equivalent where the homotopy equivalence cannot be replaced by a homeomorphism to $\mathcal{M}_0(G)$.*

Swarup's theorem does not prevent there being many homotopy equivalences between homeomorphic manifolds. A homotopy equivalence is not, in general, homotopic to a homeomorphism.

If M_1, M_2 are compact, orientable, irreducible, boundary incompressible 3-manifolds and $f : M_1 \to M_2$ is a homotopy equivalence, then f is homotopic to a (orientation preserving or reversing) homeomorphism *provided it preserves the peripheral structure*. The map f is said to preserve the peripheral structure if the isomorphism f_* of the fundamental group induced by f is such that the f_*-image of the fundamental

* There is a stronger notion of *proper homotopy equivalence* where the mappings f_1, f_2 are required to be proper: the preimage of each compact set is compact. Proper homotopy equivalence has to do with mapping the boundary of one manifold to that of the other, a much stronger requirement. This condition will not be met for the situation of Exercise 5-13.

group of each boundary component is conjugate to the fundamental group of a component of ∂M_2 [Waldhausen 1968]. If the manifold is not boundary incompressible, the statement is still true if we require instead that f restricts to a homeomorphism between the boundaries. In fact f is then homotopic to a diffeomorphism extending its boundary values; see [Waldhausen 1968; Hempel 1976, Theorem 13.6; Bonahon 2002, Theorem 3.11].

If $M_1, M_2 \cong S \times [0, 1]$ where S is a closed surface, then every homotopy equivalence $f : M_1 \to M_2$ is homotopic to a homeomorphism.

Rigidity of hyperbolic manifolds under homotopy equivalences

Recent work completed with much computer assistance has substantially extended the scope of earlier results to show that homotopy equivalences between closed manifolds have a certain topological rigidity:

Theorem 5.2.2 [Gabai et al. 2003]. *Assume $\mathcal{M}(G)$ is a closed hyperbolic manifold.*

(*i*) *If $f : M^3 \to \mathcal{M}(G)$ is a homotopy equivalence from a closed, irreducible 3-manifold M^3, then f is homotopic to a homeomorphism.*
(*ii*) *If $f, f_1 : M^3 \to \mathcal{M}(G)$ are homotopic homeomorphisms, then f_1 is isotopic to f: there is a homotopy $F(t, x) : M^3 \to \mathcal{M}(G),\ 0 \le t \le 1$ with $F(0, x) = f(x)$, $F(1, x) = f_1(x)$, such that for each t, $F(t, x)$ is a homeomorphism.*
(*iii*) *The space of hyperbolic metrics on $\mathcal{M}(G)$ is path connected.*

In short, while Mostow's Rigidity Theorem (page 157) states that hyperbolic structures are unique up to homotopy, Theorem 5.2.2 states that hyperbolic structures are unique up to isotopy.

If M^3 is itself hyperbolic, part (i) follows from Mostow's rigidity theorem. If in addition $\mathcal{M}(G)$ is Haken (Section 6.3), then (iii) follows from [Waldhausen 1968] and Mostow's theorem. If in (i) and (ii) M^3 is replaced by a not necessarily hyperbolic Haken manifold M_1^3, the statements follow from [Waldhausen 1968]. Yet the remaining cases involving non-Haken manifolds require a very deep study of the hyperbolic geometry of $\mathcal{M}(G)$.

To help understand (ii), think of a closed manifold M^3 with two hyperbolic metrics ρ_1, ρ_2 on it. By Mostow's theorem, there is a diffeomorphism $F : M^3 \to M^3$, *homotopic* to the identity, taking the geodesic α in a free homotopy class in metric ρ_2 to the geodesic α' in the same free homotopy class in metric ρ_2. Statement (ii) says that α' is not just homotopic but is isotopic to α, a subtle but nontrivial distinction! It is this that implies that the space of hyperbolic metrics on M^3 is path connected: given two hyperbolic metrics there is a diffeomorphism F such that $F^*(\rho_1) = \rho_2$.

Property (ii) implies that homotopy classes of automorphisms of a closed $\mathcal{M}(G)$ are the same as isotopy classes. Consequently the group of automorphisms of $\mathcal{M}(G)$, modulo the subgroup of those isotopic to the identity, is isomorphic to the outer automorphism group of $\pi_1(\mathcal{M}(G)) = G$. By Mostow's theorem, the outer automorphisms of $\pi_1(\mathcal{M}(G))$ are isometries.

Components of the discreteness locus

Every geometrically finite group G can be "opened up" to a group G^* so that $\mathcal{M}(G^*)$ is homeomorphic to the result of removing the interior of the solid pairing tubes from $\mathcal{M}(G)$. Jørgensen first studied this operation [1974a], but today we can apply the Hyperbolization Theorem (page 324). In other words, G^* has the property that $g \in G^*$ is parabolic if and only if g lies in a rank two parabolic subgroup — there are no rank one parabolic subgroups.

Define a *minimally parabolic group* to be one that is geometrically finite and has no rank-one parabolic subgroups. (Contrast minimally parabolic with its opposite in Theorem 4.6.3.) As the basepoint for our study, we will fix a minimally parabolic group G^*. Then (see [Anderson et al. 2000] and Theorem 5.10.12)

$$\operatorname{Int} \mathfrak{R}_{\text{disc}}(G^*) = \{H \in \mathfrak{R}_{\text{disc}}(G^*) : H \text{ is minimally parabolic}\}.$$

The Density Theorem (page 260) then implies that $\mathfrak{R}_{\text{disc}}(G^*) = \overline{\operatorname{Int} \mathfrak{R}_{\text{disc}}(G^*)}$.

Denote by $\mathcal{M}_0(G^*)$ the compact manifold resulting from removing the interiors of the cusp tori; if G^* has no rank two parabolics, $\mathcal{M}_0(G^*) = \mathcal{M}(G^*)$.

Via the compact cores, and invoking the Hyperbolization Theorem (page 324), the elements of $\mathfrak{R}_{\text{disc}}(G^*)$ are "marked" by the equivalence classes of homotopy equivalences of $\mathcal{M}_0$,

$$\mathcal{HE}(\mathcal{M}_0) = \{(M, h) : h \text{ is a homotopy equivalence } h : \mathcal{M}_0 \to M/\equiv\}.$$

Here M is another compact manifold with hyperbolizable interior. The equivalence $\equiv$ is as follows: for two homotopy equivalences of $\mathcal{M}_0$, we have $(M_1, h_1) \equiv (M_2, h_2)$ if and only if there is an *orientation preserving* homeomorphism $f : M_1 \to M_2$ such that $f \circ h_1$ is homotopic to h_2.

Within each component of $\operatorname{Int} \mathfrak{R}_{\text{disc}}(G^*)$, the marked groups are quasiconformally equivalent; two elements are the same if the equivalence f can be taken to be conformal on the boundary.

In view of the Ending Lamination Theorem (page 258), we can declare that each point of $\mathfrak{R}_{\text{disc}}(G^*)$ is determined by an element of $\mathcal{HE}(\mathcal{M}_0(G^*))$, and the ending laminations of its ends.

The case of a fuchsian closed surface group G is special in the following regard. There is a reflection, for example $j : z \mapsto \bar{z}$, that induces the *identity* automorphism of G. Thus an orientation reversing map $h : \mathcal{M}_0 \to M$ that induces an isomorphism $\theta : \pi_1(\mathcal{M}_0) \to \pi_1(M)$ can be replaced by the orientation preserving jh that also induces θ. In short, for fuchsian G, we may consider $\mathcal{HE}(\mathcal{M}_0) = \{\text{id}\}$.

The prototypical example is the *shuffle* of a rolodex or pages of a book discovered by Jim Anderson and Dick Canary [1996a]. Start with a solid torus T and its core curve c. Fix a finite system of mutually disjoint, parallel simple loops $\{\gamma_k\}$ on ∂T which are not contractible in T. Correspondingly fix a collection of surfaces $\{S_k\}$, each of some genus $g_k \geq 1$ and with a single boundary component. For greater effect assume the genera g_i are all different. Slightly thicken each S_k to obtain the compact manifolds $\{S_k \times [-\epsilon, +\epsilon]\}$. The boundary of each contains the annulus $\partial S_k \times [-\epsilon, \epsilon]$.

Attach $S_k \times [-\epsilon, \epsilon]$ by gluing $\partial S_k \times [-\epsilon, \epsilon]$ to a thin neighborhood of γ_k. The resulting manifold M is orientable and compact. By "rearranging the pages" — taking a noncyclic permutation of $\{S_k\}$, we get another manifold M_τ which is homotopy equivalent but not homeomorphic to M. The manifolds M_τ have a hyperbolic structure. When the core curve of T is removed, the corresponding hyperbolic manifolds gain a rank two cusp for which ∂T becomes a cusp torus. For more details see Exercise 5-13.

Operations akin to the shuffles just described operate in a general compact manifolds M^3 with incompressible boundary. Suppose the internal structure of M^3 contains one or more embedded solid tori T such that the components of $\partial M^3 \cap T$ are annuli each of whose central curves generate the homotopy of T; such a T is called "primitive". The above example has such a structure. A homotopy equivalence $h : {M_1}^3 \to {M_2}^3$ is called a *primitive shuffle* if there exists finite sets of primitive tori $\mathcal{T}_1 \subset {M_1}^3$ and $\mathcal{T}_2 \subset {M_2}^3$ such that h restricts to an orientation preserving homeomorphism $h : {M_1}^3 \setminus \overline{\mathcal{T}}_1 \to {M_2}^3 \setminus \overline{\mathcal{T}}_2$. For details see [Anderson et al. 2000, §2].

The following compilation of results reveal much about the structure of $\mathfrak{R}_{\mathrm{disc}}(G^*)$. As before, $\mathcal{M}_0(G^*)$ denotes the compact manifold resulting from removing any solid cusp tori from $\mathcal{M}(G^*)$.

Theorem 5.2.3 [Anderson et al. 2000; Canary and McCullough 2004]. *Assume that G^* is minimally parabolic, and $\mathcal{M}_0 = \mathcal{M}_0(G^*)$ is boundary incompressible with $\partial\mathcal{M}_0 \neq \varnothing$.*

(*i*) *The union of the closures of the components of* $\operatorname{Int}\mathfrak{R}_{\mathrm{disc}}(G^*)$ *is* $\mathfrak{R}_{\mathrm{disc}}(G^*)$ *itself — the closures cannot accumulate to an interior point.*

(*ii*) *If G^* contains no parabolics at all, then $\mathfrak{R}_{\mathrm{disc}}(G^*)$ has a finite number of components, that is, $\mathcal{HE}(\mathcal{M}_0)$ is finite.*

(*iii*) *$\mathfrak{R}_{\mathrm{disc}}(G^*)$ has infinitely many components if and only if $\mathcal{M}_0$ has the following structure: $\partial\mathcal{M}_0$ contains a torus T and there exist simple, mutually disjoint loops $\alpha_t, \alpha_1, \alpha_2 \in \partial\mathcal{M}_0$ such that* (*a*) *$\alpha_t \subset T$,* (*b*) *$\alpha_1, \alpha_2 \subset \partial\mathcal{M}_0 \setminus T$ and are not freely homotopic in $\partial\mathcal{M}_0$, but* (*c*) *α_t, α_1 and α_t, α_2 bound essential cylinders in $\mathcal{M}_0$.*

(*iv*) *The components of $\mathfrak{R}_{\mathrm{disc}}(G^*)$ are in one-to-one correspondence with the finite quotient $\widehat{\mathcal{HE}}(\mathcal{M}_0) = \mathcal{HE}(\mathcal{M}_0)/\{$primitive shuffles$\}$ that identifies two elements differing by a primitive shuffle.*

(*v*) *The closures of two components X_1, X_2 of* $\operatorname{Int}\mathfrak{R}_{\mathrm{disc}}(G^*)$ *have nonempty intersection if and only if they correspond to two elements of $\mathcal{HE}(\mathcal{M}_0)$ which are primitive shuffle equivalent, that is, if and only if they correspond to the same point in $\widehat{\mathcal{HE}(\mathcal{M}_0)}$.*

Manifolds with the properties of (ii) are constructed in Exercise 4-21. The two components X_1, X_2 of (iv) are said to *bump*. If $\zeta \in \overline{X}_1 \cap \overline{X}_2$ then every neighborhood of ζ in $\mathfrak{R}(G^*)$ intersects both X_1 and X_2. When there is bumping, $\mathfrak{R}_{\mathrm{disc}}(G^*)$ is definitely not a manifold.

Holt [2003] has shown, more generally, that if $X_1, \dots, X_n$ are components of $\operatorname{Int}\mathfrak{R}_{\rm disc}(G^*)$ such that each pair X_i, X_j bumps, then for any $K \ge 1$, there exists a geometrically finite element of $\mathfrak{R}_{\rm disc}(G^*)$ such that any K-quasiconformal deformation of it is an element of $\bigcap_i \overline{X}_i$. In particular, whenever a collection of components $\{X_i\}$ is "primitive shuffle equivalent", there exists a geometrically finite point $\zeta \in \bigcap \overline{X}_i$.

Theorem 5.2.4 [Canary and McCullough 2004]. *If the boundary $\partial\mathcal{M}_0$ is compressible, then $\mathcal{HE}(\mathcal{M}_0)$ is infinite, with the following exceptions: G is a free group, the free product of two closed surface groups, or of a closed surface group with a cyclic group, or of a cyclic group with rank-2 parabolic group.*

Using the Ending Lamination Theorem (page 258), Theorem 5.2.3 can be extended to manifolds which are not geometrically finite; see the new foreword to [Canary et al. 1987] in [Canary et al. 2006].

To see why in most cases a compressible boundary leads to an infinite number of homotopy equivalences consider the following example [Canary and McCullough 2004, p. 7]. Suppose $\mathcal{M}_0$ has compressible boundary but is not a compression body. Then it contains a geodesic α which is not homotopic into the boundary. Let $D \subset \mathcal{M}_0$ be a compressing disk and let $N \cong D \times [-\varepsilon, \varepsilon]$ be a neighborhood. Take α' with origin in $D_0 = D \times \{0\}$ to be freely homotopic to α. Cut N along D_0; drag the right side of D_0 once around α' and glue the two sides back together. A homotopy equivalence h can be constructed which is the identity outside N and inside N is the map resulting from the dragging. A loop in $\partial\mathcal{M}_0$ with nonzero intersection number with ∂D will be sent to a loop which can no longer be homotoped into $\partial\mathcal{M}_0$. Therefore h is not homotopic to a homeomorphism. If there are infinitely many free homotopy classes that contain such a curve α, then $\mathcal{HE}(\mathcal{M}_0)$ is infinite.

5.3 The quasiconformal deformation space boundary

It is an interesting fact (Exercise 5-20) that $\mathfrak{T}(G)$ is the interior of its closure $\overline{\mathfrak{T}(G)} \subset \mathfrak{R}_{\rm disc}(G)$ — it is not like a open ball with a slit to the boundary removed.

Denote the boundary of $\mathfrak{T}(G)$ by $\partial\mathfrak{T}(G)$; it is contained in $\mathfrak{R}_{\rm disc}(G)$. The study of this boundary is one of the most fascinating aspects of the subject. Every group $H \in \partial\mathfrak{T}(G)$ is the limit of an algebraically convergent sequence from $\mathfrak{T}(G)$. Therefore it is discrete and corresponds to an isomorphism $\varphi : G \to H$, by Theorem 4.1.2. Yet a boundary group is no longer a quasiconformal deformation of G; some kind of degeneration must occur as we approach it from inside $\mathfrak{T}(G)$.

A boundary group H corresponding to $\varphi \in \partial\mathfrak{T}(G)$ is called a *cusp* if it is geometrically finite. It is called a *maximal cusp* if in addition all components of $\partial\mathcal{M}(H)$ are triply punctured spheres. Thus, for a maximal cusp H, all components of $\Omega(H)$ are round disks (Exercise 2-6). The limit sets of such groups are particularly attractive, as seen in the pictures in [Mumford et al. 2002] and in Figures 5.1 and 5.7.

Here is how maximal cusps arise by *pinching* in the case that $\mathcal{M}(G)$ is a geometrically finite, acylindrical manifold. (For more general cases see the Pinching Theorem in Exercise 5-3.) Choose a maximal system $\mathcal{S}$ of mutually disjoint, simple geodesics in the hyperbolic metric on $\partial\mathcal{M}(G)$. The geodesics in $\mathcal{S}$ have the property that no two are parallel in $\partial\mathcal{M}(G)$, and of course none can be homotoped to a puncture or a point. The system $\mathcal{S}$ divides $\partial\mathcal{M}(G)$ into a union of triply connected regions — Bers coined the term *pants decomposition* because each component is homeomorphic to a pair of pants. On a surface of genus g with b punctures a pants decomposition $\mathcal{S}$ consists of $3g + b - 3$ simple loops; these divide the surface into $2g + b - 2$ pants. There are countably many homotopically different pants decompositions.

Fix an annular neighborhood A_i about each $\alpha_i \in \mathcal{S}$; choose these to be mutually disjoint. Set $B_n = \{z \in \mathbb{C} : 1/n < |z| < 1\}$. There is a quasiconformal map $f_{i,n} : A_i \to B_n$. On $\partial\mathcal{M}(G)$ take the Beltrami differential which is $(f_{i,n})_{\bar{z}}/(f_{i,n})_z$ in A_i and zero in the complement of $\bigcup A_i$. This defines a quasiconformal deformation $F_n : G \to G_n$. By Thurston's Compactness Theorem there is a subsequence $\{G_m\}$ which converges algebraically. The limit group H is a kleinian group isomorphic to G which necessarily lies on $\partial\mathfrak{T}(G)$. Furthermore H is a geometrically finite group with $\partial\mathcal{M}(H)$ homeomorphic to $\partial\mathcal{M}(G) \setminus \mathcal{S}$.

The proof involves showing that the length of the geodesics on $\{\partial\mathcal{M}(G_m)\}$ in the free homotopy classes of $\mathcal{S}$ go to zero by the pinching estimate (5.5). In the complementary components $\partial\mathcal{M}(G)\setminus\bigcup A_i$, lifted into $\Omega(G) \to \Omega(G_m)$, the conformal maps $\{F_m\}$ converge to conformal maps. Concurrently, the elements of $\{G_m\}$ corresponding to the curves of $\mathcal{S}$ converge to parabolic transformations of H.

The following result uses McMullen's technique [1991] also used to prove Theorem 5.8.4.

Theorem 5.3.1 (Cusps are dense [Canary et al. 2003; Canary and Hersonsky 2004]). *Cusps are dense on $\partial\mathfrak{T}(G)$ for any geometrically finite group G with $\partial\mathcal{M}(G) \neq \varnothing$.*

If $\partial\mathcal{M}(G)$ is connected, maximal cusps are dense on $\partial\mathfrak{T}(G)$.

In particular, maximal cusps are dense on the boundary of the deformation space of a Schottky group. However in general maximal cusps are not dense (Exercise 5-8).

Theorem 5.3.2 [Bromberg and Holt 2001]. *Suppose G is geometrically finite without rank one cusps. Assume the result $\mathcal{M}_0$ of cutting out the solid cusp tori contains an essential cylinder C such that a central loop c on C is primitive and not freely homotopic to a loop in a torus boundary component. Then there exists a point $\zeta \in \partial\mathfrak{T}(G)$ such that $U \cap \mathfrak{T}(G)$ is not connected for all small enough neighborhoods $U \subset \mathfrak{R}(G)$ of ζ. Moreover $\overline{\mathfrak{T}(G)}$ is not a manifold.*

The property of $\overline{\mathfrak{T}(G)}$ at ζ is called *self-bumping*.

For information about the primitiveness hypothesis see Exercise 4-20. The boundary point ζ will be a cusp resulting from pinching a component of ∂C. The same statement holds at the boundary of any component of $\operatorname{Int}\mathfrak{R}_{\text{disc}}(G)$. The condition on loops in C assures us that they do not determine parabolics in G. The prototype of theorems of this type is presented in Exercise 6-9, Theorem 6.6.10. Compare with Theorems 5.2.3 and 5.2.4.

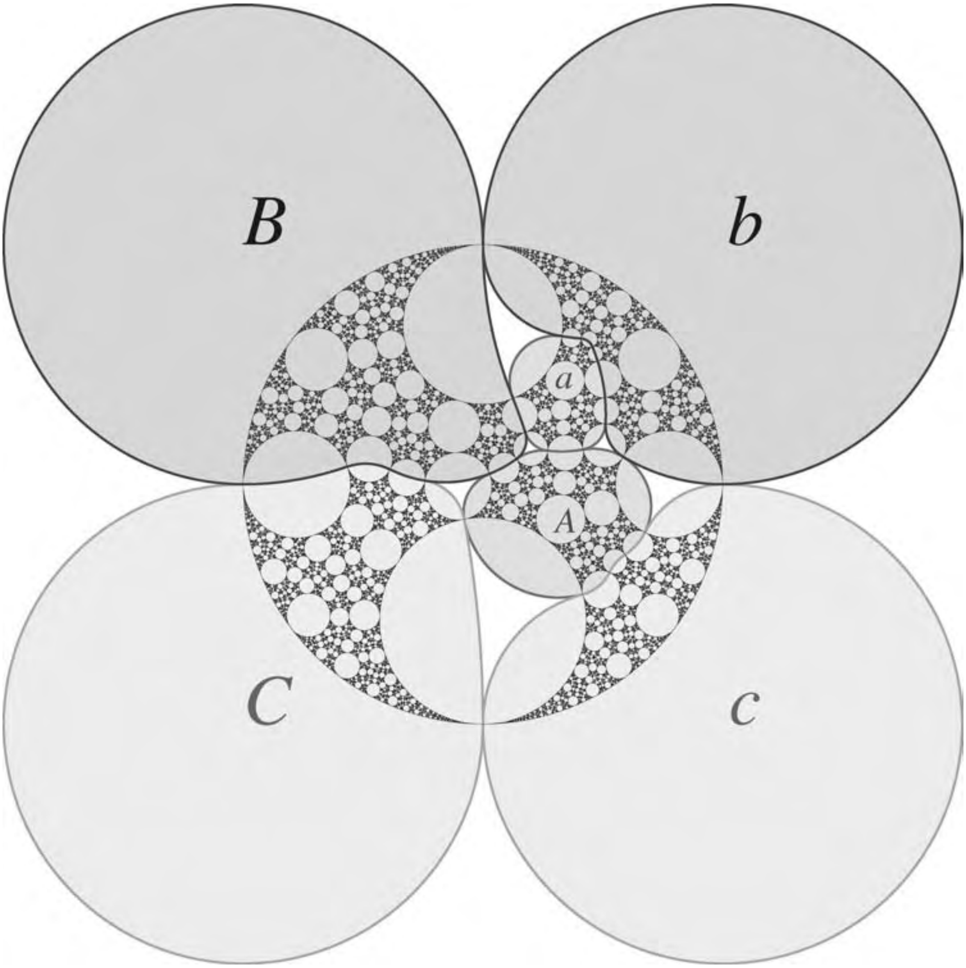


Fig. 5.1. A maximal cusp on the boundary of 3-generator Schottky space. The limiting Schottky curves are indicated.

5.4 The three great conjectures

Once geometrically finite manifolds were understood and the existence of geometrically infinite manifolds established, there emerged the daunting task of classifying the structure of these manifolds. Where are the "missing" components of $\partial\mathcal{M}(G)$? Have they left any geometric trace behind? Are the ends "wild" without a local product structure? Once the hyperbolization theorem was firmly nailed down, the following conjectures come to the fore.

The Tameness Conjecture. *The interior of every hyperbolic manifold* $\mathcal{M}(G)$ *with finitely generated* G *is homeomorphic to the interior of a compact* 3*-manifold.*

In the literature, this is called the *Marden Conjecture*. It was first raised as a question as there was little evidence for it, beyond geometrically finite groups. It is the most fundamental of the conjectures just discussed; the other two would not have been confirmed in full generality without the knowledge of tameness.

The Density Conjecture. *Every finitely generated kleinian G group is the algebraic limit of geometrically finite groups isomorphic to G.*

Groups on $\partial\mathfrak{T}(G)$ have this property. But how do we know that every group with the appropriate structure actually appears on the boundary? This enigma appeared in the context of the Bers–Maskit discovery of geometrically infinite groups on the Bers boundary. For this reason it is called the *Bers conjecture*.

The Ending Lamination Conjecture. *Hyperbolic manifolds $\mathcal{M}(G)$ with finitely generated fundamental group are completely determined by the conformal structure of $\partial\mathcal{M}(G)$ and the "ending laminations" of the geometrically infinite ends of $\mathcal{M}(G)$.*

This was proposed by Thurston when he discovered the phenomenon of ending laminations. Its solution involves an amazing bilipschitz model of hyperbolic manifolds, introduced by Minsky, that is built on the combinatorics of the closed geodesics. The proof requires the deepest analysis of the three conjectures.

It is a tribute to the power of the researchers who have entered the field in the past 20 years or so that all three conjectures have now been solved. We can now say that the main structural features of geometrically infinite hyperbolic manifolds with finitely generated fundamental groups are understood. The formal statements will come in Section 5.6.

5.5 Ends of hyperbolic manifolds

Here is a precise definition of an end, or "ideal boundary component", of an open manifold M. Exhaust M by a strictly increasing sequence of connected, compact submanifolds (with boundary) U_i:

$$\cdots \subset U_{i-1} \subset U_i \subset U_{i+1} \subset \cdots, \quad \bigcup_i U_i = M.$$

Insist that no component of $M \setminus U_i$ be compact in M (this is called a *regular exhaustion*). Correspondingly consider a nested sequence $\{V_i\}$ of open subsets of M,

$$\cdots \supset V_{i-1} \supset V_i \supset V_{i+1} \supset \cdots,$$

where each V_i is a connected component of $M \setminus U_i$. Then $\bigcap V_i = \varnothing$. The sequence $\{V_i\}$ defines an *end* or *ideal boundary component* of M.

Another such sequence $\{V'_k\}$, perhaps coming from a different exhaustion of M, defines the same end if for each large j, V_j lies in some V'_k and for each large k, V'_k lies in some V_j; an *end* is formally an equivalence class of such nested sequences. The simplest example is that the end of the manifold $\mathbb{C}$ is defined by the equivalence class of the nested sequence $V_n = \{z : |z| > n\}$.

For a hyperbolic manifold $\mathcal{M}(G)$, by an "end" we will mean an end of the interior $\mathbb{H}^3/G$. The ends are in one-to-one correspondence with the boundary components of the compact core C — this is independent of the particular choice of core C. Consequently there are only a finite number of ends (we are always assuming G is finitely generated). The appropriate complement of the compact core serves as a *neighborhood* of the end, as does any V_j that lies in it.

Thus for a finitely generated quasifuchsian group G without parabolics, $\mathcal{M}(G)$ has two ends. On the other hand if G has parabolics, $\mathcal{M}(G)$ has one end and the compact core C is a handlebody.

In the presence of parabolics the above definition of end is not the one we will normally work with for it does not account for the fact that rank one cusps correspond to what in effect are extra parts of the boundary, namely cusp cylinders. Instead we will use the refined notion of relative compact core $C = C_{\text{rel}}$ introduced in Section 3.11. There we chose $\mathcal{P}$ to be a G-invariant union of (open) horoballs at the cusps. The relative boundary in $\mathbb{H}^3/G$ of $(\mathbb{H}^3 \setminus \mathcal{P})/G$ consists of cusp tori and cusp cylinders. The boundary of the relative compact core C_{par} contains the cusp tori, and it also intersects each cusp cylinder in a closed essential annulus. Let $\mathcal{M}_p$ denote the closed submanifold $(\mathbb{H}^3 \setminus \mathcal{P})/G$ of $\mathbb{H}^3/G$.

A *relative end* E_{rel} of $\mathcal{M}(G)$ corresponds to a relative boundary component $S \subset \mathcal{M}(G) \setminus \overline{\mathcal{P}/G}$ of the relative compact core. The complementary component V of the relative compact core with relative boundary $S = \partial V$ is a neighborhood of E_{rel}. Thus for a quasifuchsian group with parabolics, $\mathcal{M}(G)$ has two relative ends.

An end E or relative end E_{rel} of $\mathcal{M}(G)$ is called *geometrically finite* if it has a neighborhood which does not intersect the *convex* core of $\mathcal{M}(G)$. Geometrically finite relative ends correspond to the components of $\partial\mathcal{M}(G)$, or equally, to the boundary components of the convex hull. It is appropriate to declare ends that correspond to cusp tori to be geometrically finite as well. If there are no parabolics, $E = E_{\text{rel}}$.

Our concern now lies with *geometrically infinite ends*: those which are not geometrically finite.

5.6 Tame manifolds

A hyperbolic manifold is called (*topologically*) *tame* if it is homeomorphic to the interior of a compact manifold.

An end E or relative end E_{rel} is called *tame* if it has a neighborhood V in $(\mathbb{H}^3 \setminus \mathcal{P})/G$ with relative boundary $\partial V = S$ such that V is homeomorphic to $S \times [0, 1)$. Here V may be taken so that S is a boundary component of a compact core or relative core.

The end E or relative end E_{rel} is said to be *compressible* if $S = \partial V$ is compressible. Otherwise, E or E_{rel} is called *incompressible*.

Each compressible end corresponds to a compressible boundary component of the compact or relative compact core which in turn corresponds to the compressible boundary component of a compression body resulting from cutting the core along incompressible surfaces (see Exercise 3-11). However note that there is more than one way to represent the end as a product. Suppose $f : S \to S$ is an orientation preserving homeomorphism, not homotopic to the identity, but extending to a homeomorphism of V which is homotopic to the identity in V. Then V is also homeomorphic to $f(S) \times [0, 1)$.

The manifold $\mathcal{M}(G)$ itself is said to be *tame* if each of its ends or relative ends is tame. Actually before the fundamental paper [Canary 1993], Thurston had worked with two notions: topologically tame (as we have defined it) and geometrically tame which requires in addition that each end be exhaustible by pleated surfaces. An explicit example of this latter phenomenon arises from manifolds fibered over the circle to be explained in Section 6.1. Canary proved that the two notions are equivalent.

From the point of view of projections from covering manifolds, geometrically infinite ends behave as if the missing surface were actually present. This is the import of the main case of the important Covering Theorem, which we have here modified by incorporating the Tameness Theorem:

Theorem 5.6.1 (Covering Theorem [Thurston 1979, Theorem 9.2.2; Canary 1996, Corollary B]). *Suppose $\mathcal{M}(H)$ has infinite volume and $G \subset H$ is a finitely generated subgroup such that the covering $\mathcal{M}(G)$ of $\mathcal{M}(H)$ has a geometrically infinite relative end $\widehat{E}_{\mathrm{rel}}$. Then there exists a neighborhood $\widehat{V} \cong S \times [0, 1)$ of $\widehat{E}_{\mathrm{rel}}$ such that the projection $\pi : \widehat{V} \to \mathcal{M}(H)$ is k-to-one for some $1 \le k < \infty$.*

This would apply for example to a geometric limit H at an algebraic limit G. The geometric limit $\mathcal{M}(H)$ has a geometrically infinite end if $\mathcal{M}(G)$ does. Even though G may have infinite index in H, each infinite end of $\mathcal{M}(G)$ behaves as a finite-sheeted cover over an end of $\mathcal{M}(H)$. This is consistent with Lemma 3.6.3.

The Tameness Theorem

Each of the following sufficient conditions of Bonahon has been a fundamental tool in dealing with the tameness question:

Bonahon's Tameness Criteria. *Either of the following conditions implies that $\mathcal{M}(G)$ is tame.*

- **A.** *G cannot be split as a free product $G = A * B$ with $A, B \neq \{\mathrm{id}\}$.*
- **B.** *In any splitting $G = A * B$ of G with $A, B \neq \{\mathrm{id}\}$ there is a parabolic $g \in G$ none of whose conjugates is contained in A or in B.*

A group G satisfying the hypothesis of Condition A is called *freely indecomposable*. Free product decompositions arise (for compact manifolds) from compression disks. Condition B says in effect that any compression disk must cut through at least one cusp cylinder. That is, while $\partial\mathcal{M}(G)$ is incompressible, the boundary of the manifold resulting after removing the solid cusp tubes is compressible. Under the splitting of G in Condition B, the two subgroups are not truly independent, rather they are tied together by a parabolic. Bonahon's conditions are satisfied in the following important cases. For the definition of compact and relative core see Section 3.11.

Corollary 5.6.2. *Suppose $\mathcal{M}(G)$ is geometrically finite and boundary incompressible. Assume there is an isomorphism $G \to H$ sending parabolics to parabolics (i.e., type preserving). Then $\mathcal{M}(H)$ is tame.*

In particular, all points on the boundary of the quasiconformal deformation space of a geometrically finite group with incompressible boundary correspond to tame manifolds.

Proof. Condition A holds for H if and only if it holds for G even though H may have new parabolics. Alternatively, it may happen that once the solid pairing tubes are removed from $\mathcal{M}(G)$ the resulting manifold is no longer boundary incompressible. In these cases Condition B comes into play, and the condition will be maintained in H. A boundary point of a quasiconformal deformation space $\mathfrak{T}(G)$ is the algebraic limit $\theta : G \to H$ of a sequence of isomorphisms $\{\theta_n : G \to G_n\}$ to geometrically finite groups in the space. □

There had been a steady advance in understanding tameness before the complete answer was found. Partial results had been obtained by Thurston, Brock, Bromberg, Canary, Evans, Minsky, Ohshika, individually and in collaboration; for example see [Brock et al. 2003; Ohshika 2005]. Here we record two of the notable results (compare with Theorem 4.6.2).

Theorem 5.6.3 [Canary and Minsky 1996; Evans 2004b]. *Suppose $\{\theta_n : G \to G_n\}$ is a type preserving sequence of isomorphisms where each $\mathcal{M}(G_n)$ is known to be tame. Suppose the sequence converges algebraically and geometrically to a group H which has no new parabolics. Then H is tame as well.*

Theorem 5.6.4 [Brock and Souto 2006]. *The algebraic limit of any sequence of geometrically finite groups is tame.*

This implies:

Corollary 5.6.5. *If the density conjecture is true, all hyperbolic manifolds $\mathcal{M}(G)$ with finitely generated G are tame.*

The resolution of the tameness conjecture was first announced by Ian Agol. A different proof was being developed independently by Danny Calegari and David Gabai. Both sets of authors credit discussions with Mike Freedman.

The Tameness Theorem [Agol 2004; Calegari and Gabai 2004]. *Every hyperbolic manifold $\mathcal{M}(G)$ with finitely generated G is tame.*

In view of Bonahon's Condition A, it suffices to restrict consideration to the compressible ends.

Agol's proof makes heavy use of manifolds of pinched negative curvature. In particular this allows him to remove the rank one and rank two cusps. He then uses Canary's trick of finding a "diskbusting" curve to construct a two-sheeted cover which also has finitely generated fundamental group but a given end is now incompressible. More of the details are outlined in Exercises 5-17 and 5-18.

The Calegari–Gabai proof is instead centered on the existence of "shrinkwrapped surfaces". Namely, suppose Γ is a finite collection of mutually disjoint, simple closed geodesics in $\mathcal{M}(G)$ where G has no parabolics. Let S be a closed, incompressible surface in $M = \mathcal{M}(G) \setminus \Gamma$. Then S can be *shrinkwrapped* in $\overline{M}$: there is an isotopy $F : S \times [0, 1] \to \overline{M}$ with $F(\cdot\,, 0) = S$, $F(\cdot\,, t)$ is an embedding of S in M, $0 \leq t \leq 1$, and $T = F(\cdot\,, 1)$ is a CAT(-1) surface (this is a hyperbolic-like metric property; see [Ohshika 2002]), which is a minimum for hyperbolic area among all surfaces in the homotopy class. The minimizers are likely to abut upon Γ so the actual structure may be more complicated, although it will remain a CAT(-1). Using a sequence of geodesics which exit a geometrically infinite end, the authors find a sequence of shrinkwrapped surfaces trapped between the successive geodesics. By establishing a uniform bound on their diameters, they show the shrinkwrapped surfaces exit the end. Recently their proof has been simplified in [Soma 2005].

Earlier it was shown in [Canary 1993] that for an end to be tame there must exist a neighborhood V of the end, and a sequence of (not necessarily embedded) pleated surfaces in V, each homotopic within V to $\partial V = S$, that exit the end or relative end. This has turned out to be an important ingredient in the tameness proofs.

Souto [2005] had proved that $\mathcal{M}(G)$ is tame if its interior can be exhausted by a nested union of compact cores. It is not enough to find a sequence of mutually disjoint surfaces $\{S_n\}$ of the same topological type exiting each end. It is necessary to know that each pair (S_n, S_{n+1}) bounds a region homeomorphic to $S_n \times [0, 1]$; that is, S_{n+1} is homotopic to S_n in such a way that given any compact set, the homotopy does not meet it for all large indices. It would be nice to be able to study the ends using only a fundamental polyhedron!

As an example, if G is a free group of rank two, $\operatorname{Int} \mathcal{M}(G)$ is homeomorphic to the interior of a handlebody, even if $\partial \mathcal{M}(G) = \varnothing$!

Building on earlier results of Thurston and Bonahon, Canary [1993] proved the forty-year old Ahlfors' Conjecture for tame manifolds. Canary's result, coupled with Theorem 5.6.4, guarantees that any algebraic limit of geometrically finite groups satisfies Ahlfors' conjecture. Ahlfors himself had proved it for geometrically finite manifolds. A prior special case was treated in [Ohshika 2005]. Adding in the Tameness Theorem, we now have:

Theorem 5.6.6 (Ahlfors Conjecture). *For any finitely generated group G, either $\Lambda(G) = \mathbb{S}^2$, or $\Lambda(G)$ has 2-dimensional Lebesgue measure zero. Moreover if $\Lambda(G) = \mathbb{S}^2$ the action of G is ergodic: There does not exist a pair of disjoint G-invariant sets each of positive measure.*

The Ending Lamination Theorem

It is easier to characterize ending laminations in the case of an incompressible relative end E_{rel} of $\mathcal{M}(G)$ — if there are no parabolics than E_{rel} is just an end E. Since we now know that infinite ends are tame, it has a neighborhood V (with $\partial V = S$) homeomorphic to $S \times [0, 1)$ with injective inclusion $\phi : \pi_1(S) \to G$. We can take

S as an embedded pleated surface in $\mathcal{M}(G)$. The surface S can be represented as a finite area surface group $S^* = \mathbb{H}^2/\Gamma$ whose punctures correspond to simple loops on cusp cylinders. However it is still possible that a loxodromic in Γ corresponds to a parabolic in G.

The isomorphism $\phi^* : \Gamma \to \pi_1(S) \hookrightarrow G$ determines a one-to-one map from the ordered fixed point pairs $(u_a, u_r) \in \partial\mathbb{H}^2$ of the loxodromic elements $\gamma \in \Gamma$ to the fixed point pairs (u'_a, u'_r) of the $\phi(\gamma)$ in the limit set $\Lambda(G)$, provided $\phi(\gamma)$ is loxodromic. The image of a sequence of loxodromic axes which converge to a geodesic in $\mathbb{H}^2$ will converge to a geodesic in $\mathbb{H}^3$ oraa point in $\mathbb{H}^3$.

A geodesic lamination $\Lambda \subset \mathbb{H}^2/\Gamma = S^*$ is said to be *realizable* in $\mathcal{M}(G)$ if there is a pleated surface $f : S^* \to S \subset \mathcal{M}(G)$ whose bending lamination contains $f(\Lambda)$ [Canary et al. 1987, Theorem 5.3.9]. In particular, each leaf corresponds to a geodesic in $\mathcal{M}(G)$. A measured lamination is said to be realizable in $\mathcal{M}(G)$ if and only if its support is realizable.

For compressible ends, lots of geodesics on S are not realizable — the compressible ones for example. Only in the case that E_{rel} is geometrically finite and incompressible are all laminations of S realizable.

Suppose $\Lambda = \lim \gamma_k$ where $\gamma_k \subset S^*$ are simple geodesics. Then $\phi(\Lambda) = \lim \phi(\gamma_k)$ exists as a lamination in $\mathcal{M}(G)$ if and only if the sequence $\{\phi(\gamma_k)\}$ lies in a compact subset of $\mathcal{M}(G)$.

The simplest example of a nonrealizable lamination occurs when V is a neighborhood of a pinched surface in $\partial\mathcal{M}(G)$. Then the geodesic on S that represents the pinching is not realizable. Of course we have excluded this by use of the relative ends.

Here is a useful estimate. Suppose $\gamma \subset \mathcal{M}(G)$ is a closed geodesic, and $\gamma^* \subset \mathcal{M}(G)$ is a simple loop freely homotopic to γ such that the closest distance of γ^* to γ is r. Then

$$\mathrm{Len}(\gamma) \le \frac{\mathrm{Len}(\gamma^*)}{\cosh r}. \tag{5.1}$$

Proof. Consider the tubular neighborhood C_r of radius r about γ. The shortest simple loop on ∂C_r freely homotopic to γ has length $\cosh r \,\mathrm{Len}(\gamma)$ (see Exercise 1-4). □

Suppose S is an incompressible pleated surface inside $\mathcal{M}(G)$ that faces a geometrically infinite, incompressible relative end $E_{\text{rel}} \cong S \times [0, 1)$. A sequence of geodesics in E_{rel} is said to *exit* E_{rel} if, given any compact subset $K = S \times [0, r] \subset V$, at most a finite number of elements of the sequence have nonempty intersection with K.

Lemma 5.6.7 [Thurston 1979, §9.3; Bonahon 1986]. (*i*) *Assume $\{\gamma_n\}$ is a sequence of closed geodesics exiting E or E_{rel}. Assume each γ_n is freely homotopic within U to a simple loop $\gamma_n^* \subset S$ which is a geodesic in the hyperbolic metric on S. Then* $\mathrm{Len}_S(\gamma_n^*) \to \infty$.

(*ii*) *Suppose α, β are closed geodesics freely homotopic within U to simple geodesics $\alpha^*, \beta^* \subset S$, and are both of distance exceeding r from S. Given $\epsilon > 0$, assume that each of α, β is either disjoint from the ϵ-thin part of $\mathcal{M}(G)$ or is the core* *of an* ϵ*-tubular neighborhood about itself. Then on* S*, there exists a constant* $C = C(\epsilon)$

$$\iota(\alpha^*, \beta^*) \le Ce^{-r}\, \mathrm{Len}_S(\alpha^*)\, \mathrm{Len}_S(\beta^*) + 2. \tag{5.2}$$

(*iii*) *Suppose* $\{\alpha_n\}, \{\beta_n\}$ *are exiting sequences such that their realizations in* S *converge to measured laminations:* $\alpha_n^*/c_n \to (\Lambda_1, \mu)$, $\beta_n^*/d_n \to (\Lambda_2, \nu)$. *Then* $\iota(\mu, \nu) = 0$, *so that no leaf of* Λ_1 *crosses a leaf of* Λ_2 *and* $\Lambda_1 \cup \Lambda_2$ *is also a lamination; possibly* $\Lambda_1 = \Lambda_2$.

Outline of proof. Statement (i) has an elementary proof depending only on the fact that S has finite topological type. For if, for an infinite subsequence $\mathrm{Len}_S(\gamma_k^*) < M < \infty$ then for all except at most a finite number of indices, the geodesics $\{\gamma_k^*\}$ would coincide with a fixed closed geodesic γ^*. The corresponding $\{\gamma_k\}$ would also coincide.

The proof of (ii) is based on a detailed study of the interaction of the free homotopy cylinders between each pair of geodesics (α, α^*), (β, β^*). The cylinders can be assumed to be transverse to each other (if $\iota(\alpha^*, \beta^*) = 0$, (ii) is vacuous). Also used is Equation (8.32).

Property (iii) follows directly from Equation (5.2). That is why (5.2) plays a key role in the existence theory. Compare with Exercise 3-35. □

The following theorem was initiated by Thurston [1979], then filled out by Bonahon [1986] and Canary [1993]. A good overview can be found in [Minsky 1994a]. It was originally proved under the assumption of tameness; here we complete their theorem by taking account of that. Canary's paper dealt with compressible ends, reducing them to incompressible ends as described in Exercise 5-17, and clearing the path to their analysis.

Theorem 5.6.8 (Existence of ending laminations). (*i*) *Suppose* E_{rel} *is an incompressible relative end with neighborhood* $V = E_{\mathrm{rel}} \subset \mathcal{M}(G) \cong S \times [0, 1)$, *where* $S = \partial V$ *is a finite area pleated surface. The end* E_{rel} *is geometrically infinite if and only if there is a sequence of closed geodesics* $\{\gamma_n\}$ *which exit* E_{rel} *with each* γ_n *freely homotopic in* V *to a simple geodesic* $\gamma_{S,n} \subset S$.

There is a uniquely determined measurable lamination $\Lambda(E_{\mathrm{rel}}) \subset S$ *such that* $\Lambda(E_{\mathrm{rel}}) = \lim \gamma_{S,n}$, *for any such exiting sequence.*

(*ii*) *Suppose instead that* E_{rel} *is a compressible end with a neighborhood*

$$V \subset \mathcal{M}(G) \cong S \times [0, 1),$$

where $S = \partial V$ *is a finite area pleated surface. It is geometrically infinite if and only if there is an sequence of closed geodesics* $\{\gamma_n\}$ *which exit* E_{rel} *with each* γ_n *freely homotopic in* V *to an incompressible simple geodesic* $\gamma_{S,n} \subset S$.

Once the product structure of the neighborhood V *is fixed, there is a uniquely determined measurable lamination* $\Lambda(E_{\mathrm{rel}}) \subset S$ *such that* $\Lambda(E_{\mathrm{rel}}) = \lim \gamma_{S,n}$ *for any such exiting sequence.*

It needs to be emphasized that in the compressible case, there are countably many possibilities for expressing the end as a product. Without a specific specification, then $\Lambda(E_{\text{rel}})$ is determined only up to the action by the group of automorphisms of S which are homotopic to the identity in $\mathcal{M}(G)$ — but not in S itself. Here S may be taken as the boundary component of a compact core that faces E_{rel}.

The lamination $\Lambda(E_{\text{rel}})$ is called the *ending lamination* of the end E_{rel}. It is constructed in [Thurston 1979, Theorem 9.3.2] in terms of limits of measured laminations, supported by simple geodesics in S, for which the corresponding geodesics in $\mathcal{M}(G)$ exit E_{rel}. The key point is that if $\Lambda_1, \Lambda_2 \subset S$ are two limiting laminations, as a consequence of Lemma 5.6.7 their leaves do not cross. Therefore $\Lambda(E_{\text{rel}})$ can be defined as the union of all the limits.

An ending lamination does not come with it any particular measure and may support projectively a number of measures, or may be uniquely ergodic and support only one.

The ending lamination $\Lambda(E_{\text{rel}})$ of an incompressible relative end is a *maximal, arational* geodesic measurable lamination. In particular:

- Each simple geodesic on S is transverse to $\Lambda(E_{\text{rel}})$. In fact if ν is any measured lamination on S with support different than $\Lambda(E_{\text{rel}})$, then the geometric intersection number (Section 3.9) satisfies $\iota(\mu, \nu) \neq 0$, where μ is a measure on $\Lambda(E_{\text{rel}})$.
- Each half-leaf in $\Lambda(E_{\text{rel}})$ is dense in $\Lambda(E_{\text{rel}})$.
- $\Lambda(E_{\text{rel}})$ is not a proper sublamination of any measurable lamination; each component of $S \setminus \Lambda(E_{\text{rel}})$ is an ideal polygon, possibly containing a puncture. (A geodesic that divides an ideal polygon is isolated and hence cannot support a measure.)

The recently announced blockbuster proof of the Ending Lamination Conjecture is built on the pioneering work of Yair Minsky [2003a] in constructing a Lipschitz model of $\mathcal{M}(G)$. The model is organized around the set of simple closed geodesics and its combinatorics as dictated by the curve complex on surfaces representing the ends with its metric (Exercise 5-15), as worked out jointly with Masur [1999; 2000]. The proof was completed by the team of Brock, Canary, and Minsky [Brock et al. 2004, $\geq$ 2007]. They completed the Minsky–Masur work by showing there is a bilipschitz map $\Phi : \text{Model} \to \operatorname{Int} \mathcal{M}(G)$ depending essentially on the ending laminations (exiting geodesics). The *coup de grâce* was delivered by the Tameness Theorem [Agol 2004; Calegari and Gabai 2004].

Minsky [2003b] describes the combinatorial basis of the proof. Earlier [1999] he had solved the ending lamination conjecture for the once-punctured torus case — which satisfies Bonahon's Condition B.

Ending Lamination Theorem. *Suppose $\phi : G_1 \to G_2$ is isomorphism between finitely generated groups so that $\phi(g)$ is parabolic if and only if g is so. Assume ϕ is induced by a homeomorphism $\Phi : \mathcal{M}(G_1) \to \mathcal{M}(G_2)$ such that $\Phi : \partial\mathcal{M}(G_1) \to \partial\mathcal{M}(G_2)$*

is homotopic to a conformal mapping, and that corresponding geometrically infinite ends E_{rel} and $\Phi(E_{\mathrm{rel}})$ have the same ending laminations. Then $\phi : G_1 \to G_2$ is realized by a conjugation.

In short, a hyperbolic manifold is determined uniquely up to isometry by its topological structure and its ending invariants.

A proof that all possibilities for ending laminations actually occur appears (for the parabolic free case) in [Ohshika 2003], using [Kleineidam and Souto 2002]. See also the Double Limit Theorem (page 314).

The end invariants are not continuous in $\mathfrak{R}(G)$. For a simple example consider a cusp on $\partial\mathfrak{R}_{\mathrm{disc}}(G)$. As we make a "tangential" approach to the cusp from within the deformation space, or approach the cusp along its boundary, the geometric limit (as we may assume) is larger than the algebraic; the limit sets of the approximates will look more and more like the limit set of the geometric limit. For a full discussion see [Brock 2000; 2001a].

The following assertion follows from Theorem 5.6.8 and Formula (5.1); see also Exercise 5-22.

Corollary 5.6.9. *Suppose $\{G_n\}$ is a sequence of quasifuchsian groups converging algebraically and geometrically to a singly degenerate group H. Let (Λ_n, β_n) denote the bending lamination of the boundary component C_n of the convex core of $\mathcal{M}(G_n)$ that is approaching the infinite end E of H. Then $\{(\Lambda_n, \beta_n)\}$ converges to a measured lamination (Λ, β) in the reference surface S such that Λ is not realizable in $\mathcal{M}(H)$. That is, Λ is the ending lamination of E.*

When applied more generally to algebraic and geometric limits of geometrically finite manifolds, this gives a "natural" way of finding the ending laminations.

Alternately, instead of bringing in the bending laminations, the process can be described in terms of the hyperbolic metrics g_n on the degenerating component or components of $\partial\mathcal{M}(G_n)$. These metrics are obtained from pulling over the metric in $\mathbb{H}^2$ to the degenerating component or components of $\Omega(G_n)$ by the Riemann maps. The one or two sequences $\{g_n\}$ converge to measured laminations whose support(s) are the ending lamination(s) in the sense of Thurston (page 280); see the Double Limit Theorem (page 314). Theorem 3.11.2 displays how the geometry of the convex core boundary and the geometry of the surface facing it "at infinity" are related.

The Ending Lamination Theorem has the following consequence proved in [Brock et al. 2004] under the assumption of tameness; a prior version (including compressible ends) was given by [Ohshika 1998b]. The second statement is an application of Sullivan's Theorem (page 158). It was first proved in [Brock et al. 2004] for tame manifolds. Once again we have applied tameness to complete the original statements.

Theorem 5.6.10 (Topological rigidity). *Suppose $\mathcal{M}(G)$, $\mathcal{M}(H)$ have incompressible ends. Assume there is an orientation preserving homeomorphism $\Psi : \mathbb{S}^2 \to \mathbb{S}^2$ which induces an isomorphism $\psi : G \to H$ so that $\Psi \circ g(z) = \psi(g) \circ \Psi(z)$ for all $z \in \mathbb{S}^2$ and*

$g \in G$. Then there is a quasiconformal mapping $F : \mathbb{S}^2 \to \mathbb{S}^2$ that likewise satisfies $F \circ g(z) = \psi(g) \circ F(z)$ for all $z \in \mathbb{S}^2$ and all $g \in G$.

If Ψ is conformal on $\Omega(G)$, or if $\Omega(G) = \varnothing$, then F is a Möbius transformation.

The essence of the proof is to show $\mathcal{M}(G)$ and $\mathcal{M}(H)$ have the same ending laminations.

Before the Ending Lamination Theorem was announced, the ending lamination conjecture was proved in [Minsky 2001] under the following assumption: There exists a positive lower bound $\delta > 0$ for the length of all closed geodesics in the interior of the manifold. This condition forces a certain uniformity in how pleated surfaces converge to the ends. Yet in general for geometrically infinite manifolds, in fact at a dense set of points in the boundary of the deformation spaces of any geometrically finite group, the uniform lower bound condition is not satisfied [McMullen 1991, Corollary 1.6; Canary et al. 2003; Canary and Hersonsky 2004].

A manifold $\mathcal{M}(G)$ satisfying the uniform lower bound condition is said to have *bounded geometry*. Geometrically finite manifolds automatically have bounded geometry. For a geometrically infinite example, see §6.1.3.

The Density Theorem

The solution of the thirty-year old density conjecture in the tame case was proved by Ohshika [1990; 2004], assuming the Tameness and Ending Lamination Theorems. Density for manifolds with incompressible ends without cusps was proved using cone manifold techniques in [Bromberg 2004; Brock and Bromberg 2003; 2004]. Also pertinent in the compressible case are [Kleineidam and Souto 2002; 2003; Lecuire 2004c]. These results show that every ending lamination of a kleinian manifold can be achieved by an algebraically convergent sequence of geometrically finite groups. The Double Limit Theorem (page 314) was the first of this type of result. Also see [Brock et al. 2004; $\geq$ 2007].

Density Theorem. *Every finitely generated kleinian group is the algebraic limit of geometrically finite groups.*

In more detail, suppose G is geometrically finite and $\theta : G \to H$ is an isomorphism to a group $H \in \mathfrak{R}_{\text{disc}}(G)$. There exists a sequence of isomorphisms $\{\theta_n : G \to H_n \in \mathrm{Int}\, \mathfrak{R}_{\text{disc}}(G)\}$ to geometrically finite groups which converges algebraically to θ. That is, the closure $\overline{\mathrm{Int}\, \mathfrak{R}_{\text{disc}}}$ equals $\mathfrak{R}_{\text{disc}}(G)$.

Recall from Theorem 4.6.3 that the sequence can be chosen so that $\theta_n\theta^{-1}$ also preserves parabolics, but then $\theta_n(G)$ will not be in the same deformation space if H has new parabolics.

Density implies tameness by Corollary 5.6.5. That tameness implies density is a consequence of the Ending Lamination Theorem, as described above. Thus we have a complete answer to what was suspected earlier:

Corollary 5.6.11. *The tameness conjecture holds if and only if the density conjecture holds.*

Untame manifolds

Lest one think that tameness is self-evident, it is worth pondering an example of Fox and Artin [1948]: There exists a wild embedding of $\mathbb{S}^2$ into $\mathbb{S}^3$ such that both complementary components are simply connected, and neither component is homeomorphic to the (open) 3-ball $\mathbb{B}^3$. That σ is a wild embedding means that there is no homomorphism of S^3 that carries σ to S^2. In contrast, wild embeddings of $\mathbb{S}^1$ in $\mathbb{S}^2$ do not exist; the closure of each complementary component of a simple closed curve is homeomorphic to the closed ball.

The same example of Fox and Artin shows that there exists a polyhedral plane in $\mathbb{R}^3$ such that the closure of neither complementary region is homeomorphic to a closed half-space $\mathbb{R}^2 \times [0, \infty)$. We are grateful to Tom Tucker for the reassurance that this awful possibility does not arise in our analysis:

Theorem 5.6.12 [Tucker 1975]. *Suppose $S \subset \mathcal{M}(G)$ is an incompressible surface embedded in Int $\mathcal{M}(G)$ with nontrivial fundamental group. Let S^* be a component of $\{\pi^{-1}(S)\}$ in $\mathbb{H}^3$. Then the closure of each complementary component of S^* in $\mathbb{H}^3$ is homeomorphic to a closed half-space.*

A cornucopia of other weird examples are presented in [Scott and Tucker 1989], including one of Peter Scott already mentioned in [Marden 1974b]: Given a closed surface S of genus ≥ 1 there is a 3-manifold M^3 with the properties (i) $\partial M^3 = S$, (ii) the injection $\pi_1(S) \hookrightarrow \pi_1(M^3)$ is an isomorphism, (iii) the universal cover of M^3 is the closed upper half-space H, (iv) for any deck transformation T, $H/\langle T\rangle \cong (\mathbb{S}^1 \times \mathbb{R}) \times [0, 1)$, but (v) $M^3 \neq S \times [0, 1)$! Singly degenerate groups have the first four properties.

Tucker [1974] gave the following example. Let T_0' be a solid torus $\mathbb{D} \times \mathbb{S}^1$. Embed T_0' in a larger solid torus $T_0 \subset T_1$ so that T_0 is knotted, and homotopy equivalent to T_1. Take an infinite sequence of nested solid tori $T_2 \subset T_3 \subset T_4 \subset \dots$ so that there is a homeomorphism of $T_{k+1} \setminus T_k$ onto $T_1 \setminus T_0$. Set $M = \bigcup_k T_k$. Then $\pi_1(M)$ is infinite cyclic and covered by $\mathbb{R}^3$, but M cannot be embedded in any compact manifold. Also $\pi_1(M \setminus T_0)$ is not finitely generated.

A necessary condition that a noncompact irreducible manifold M^3 be homeomorphic to the interior of a compact manifold is that for every compact submanifold $K \subset M^3$, $\pi_1(M^3 \setminus K)$ is finitely generated. It is not so far from being sufficient; see [Tucker 1974].

5.7 Quasifuchsian spaces

We will start this section with a fuchsian group G acting in the upper and lower half-planes such that $R = \text{LHP}/G$ is a surface of genus g with $b \geq 0$ punctures satisfying $3g + b - 3 > 0$. The reflected surface $R' = \text{UHP}/G$ is anticonformally equivalent to R under reflection $J_0(z) = \bar{z}$ in $\mathbb{R}$.

The simplest deformation spaces are the quasifuchsian spaces $\mathfrak{T}(G)$. By the principle of simultaneous uniformization (Section 3.8), the points of this space can be

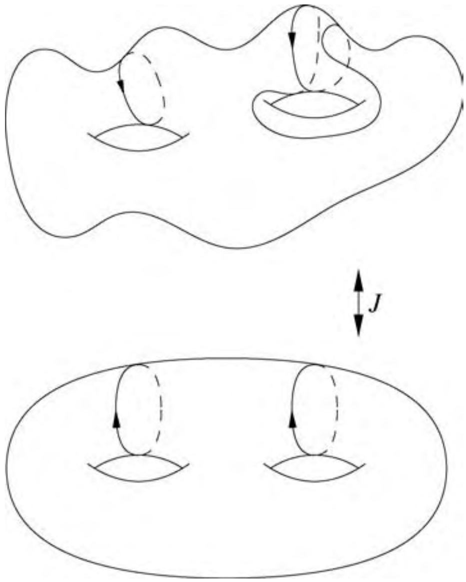


Fig. 5.2. A genus-2 quasifuchsian group. J is an orientation reversing homeomorphism pairing the curves.

described by the triples

$$\mathfrak{T}(G) = \{(S_{\text{bot}}, S^{\text{top}}; J), F\}.$$

Here S_{bot}, S^{top} are Riemann surfaces quasiconformally equivalent to R, R' respectively and J is an orientation reversing involution $S_{\text{bot}} \leftrightarrow S^{\text{top}}$. The marking is fixed by the homotopy class $[F]$ of a quasiconformal mapping $R \to S_{\text{bot}}$. Each triple is associated with a quasifuchsian group H with $\mathcal{M}(H) \cong S_{\text{bot}} \times [0, 1]$ in the following manner:

(i) H itself is uniquely determined by the triple $(S_{\text{bot}}, S^{\text{top}}; J)$, up to conjugation.
(ii) $\partial\mathcal{M}(H) = S_{\text{bot}} \cup S^{\text{top}}$ and J extends to an orientation reversing, fiber preserving involution $J : \mathcal{M}(H) \to \mathcal{M}(H)$.
(iii) There exists a quasiconformal map $F : \mathcal{M}(G) \to \mathcal{M}(H)$ such that $F(R) = S_{\text{bot}}$, $F(R') = S^{\text{top}}$ and $F J_0 F^{-1}$ is homotopic to J.
(iv) F is the projection of the restriction to $\Omega(G) = \text{LHP} \cup \text{UHP}$ of a quasiconformal map $F^* : \mathbb{S}^2 \to \mathbb{S}^2$. F^* determines an isomorphism $\varphi : G \to H$.

We will write $\partial_{\text{bot}}\mathcal{M}(H) = S_{\text{bot}}$ and $\partial^{\text{top}}\mathcal{M}(H) = S^{\text{top}}$.

Bers slices

The *Bers slice* $\mathcal{B}(R) \equiv \mathcal{B}(G) \subset \mathfrak{T}(G)$ determined by the Riemann surface $R = \text{LHP}/G$ is defined as the subset of triples

$$\mathcal{B}(R) = \{\big((S_{\text{bot}}, S^{\text{top}}; J), F\big) \in \mathfrak{T}(G) \big| F : R \to S_{\text{bot}} \text{ is conformal}\}.$$

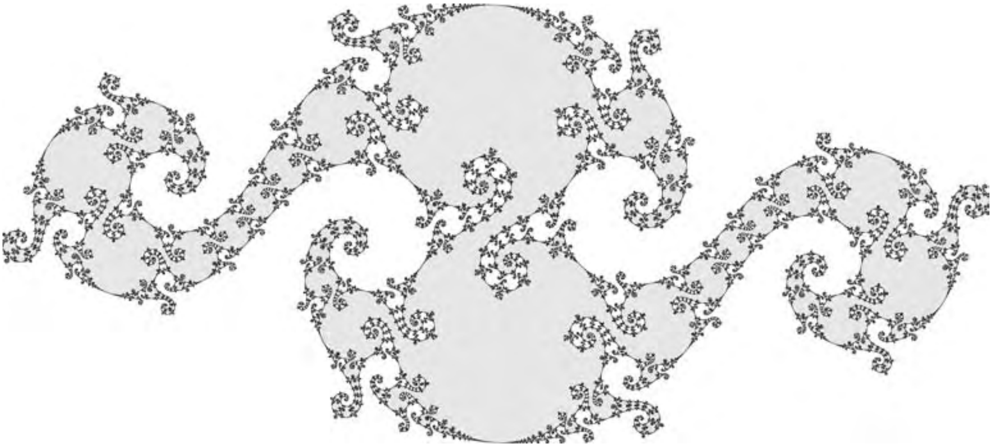


Fig. 5.3. A once-punctured torus quasifuchsian group near the boundary of the quasifuchsian space.

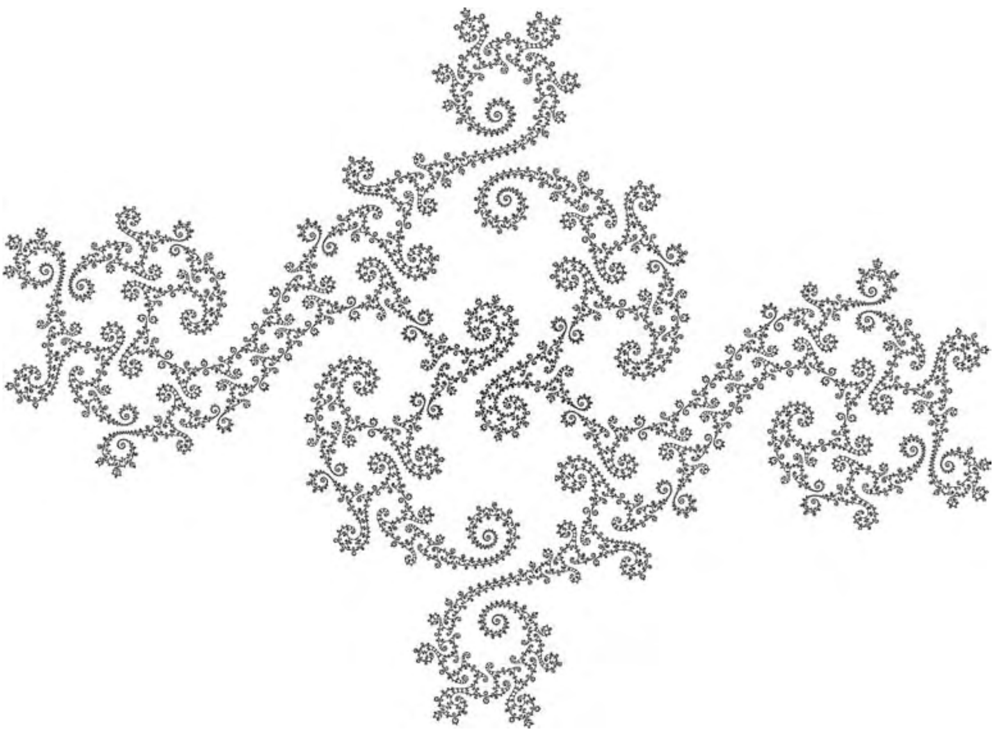


Fig. 5.4. Slightly opening the cusp of Figure 5.3 results in this Schottky group. The limit set is totally disconnected but very close to a quasicircle.

In the deformations, the bottom surface remains conformally equivalent to R; the mapping F is quasiconformal on R' and *conformal* on R. A lifted map F^* is a Riemann map of the LHP onto the component $F^*(\text{LHP}) = \Omega_{\text{bot}}$ of $\Omega(H)$. Of course F^* satisfies the relations $F^* \circ g(z) = \varphi(g) \circ F^*(z)$ for all $g \in G,\ z \in \text{LHP}$. So the image group H is determined, up to normalization, by the conformal mapping F^* on LHP.

A Bers slice is perhaps the most useful realization of the Teichmüller space (Section 2.8) of the Riemann surface R'; S^{top} runs through all possible quasiconformal deformations of R', all the while S_{bot} remaining conformally fixed as R. The initial discovery by Bers thrilled all in the field because in its realization as a slice within a space of complex matrices, the complex structure on Teich(R) becomes more accessible [Ahlfors 1966].

The slice $\mathcal{B}(R)$ is a complex analytic manifold of $\mathbb{C}$-dimension $3g+b-3$. It is also a metric space in the Teichmüller metric (page 81). The whole quasifuchsian space $\mathfrak{T}(G)$ as a holomorphic submanifold of the representation variety has twice the dimension.

At least in the case that R is a *closed* surface, there is an extension $\mathcal{B}^*(R)$ of the Bers slice, called the *extended Bers slice*, which is a *properly embedded* submanifold of $\mathfrak{R}(G)$ [Gallo et al. 2000, Theorem 11.4.1]. It is the space of projective structures on the Riemann surface R (see Exercise 6-8).

Each choice of complex structure on R determines a different Bers slice.

Masaaki Wada's program OPTi [≥ 2007b; 2006] allows the interactive exploration of the limit sets of quasifuchsian once-punctured torus groups. The groups are parametrized by Jørgensen complex probabilities Exercise 1-34.

The team of Y. Komori, T. Sugawa, M. Wada, and Y. Yamashita was the first to succeed in visualizing the complex 1-dimensional Bers slice of once-punctured torus space — based on the square torus. The method is explained in [Komori and Sugawa 2004]. Their work has recently been augmented by studies of David Dumas [2004], who analyzes the slice as an island in the archipelago which is the discreteness locus of $\mathcal{B}^*(R)$, lying in the sea of indiscreteness. Dumas' picture is the frontispiece of this book, with a closeup of the Bers slice in Figure 5.5 on the next page. This shows the Bers slice in the once-punctured torus space that is based on the hexagonal torus. The combinatorics of the Ford polyhedron give a tiling of the quasifuchsian space (Jørgensen). This tiling, restricted to the slice, gives the tiling by triangular regions you see here — this is a special property of the hexagonal slice. The boundary of the slice is a Jordan curve [Minsky 1999] and the cusps you see are dense on it [McMullen 1991]. These are indeed geometric cusps [Miyachi 2003].

There is a natural action of the Teichmüller modular group or mapping class group (see page 81 and Exercise 5-6) by isometries (and biholomorphic mappings) on $\mathcal{B}(R)$ akin to the action of a fuchsian group on $\mathbb{H}^2$. If τ is an orientation preserving automorphism of the surface $R = S_{\mathrm{bot}}$, realized as a quasiconformal automorphism if there are punctures, then τ induces the following action:

$$\tau : (S_{\mathrm{bot}}, S^{\mathrm{top}}, J) \mapsto (S_{\mathrm{bot}}, S^{\mathrm{top}}, J \circ \tau).$$

In this action the conformal types of the surfaces do not change, what changes is the *topological* relationship between the bottom and the top. The action of the iterates $\{\tau^n\}$ is described in Theorem 5.9.1 and Exercise 5-11.

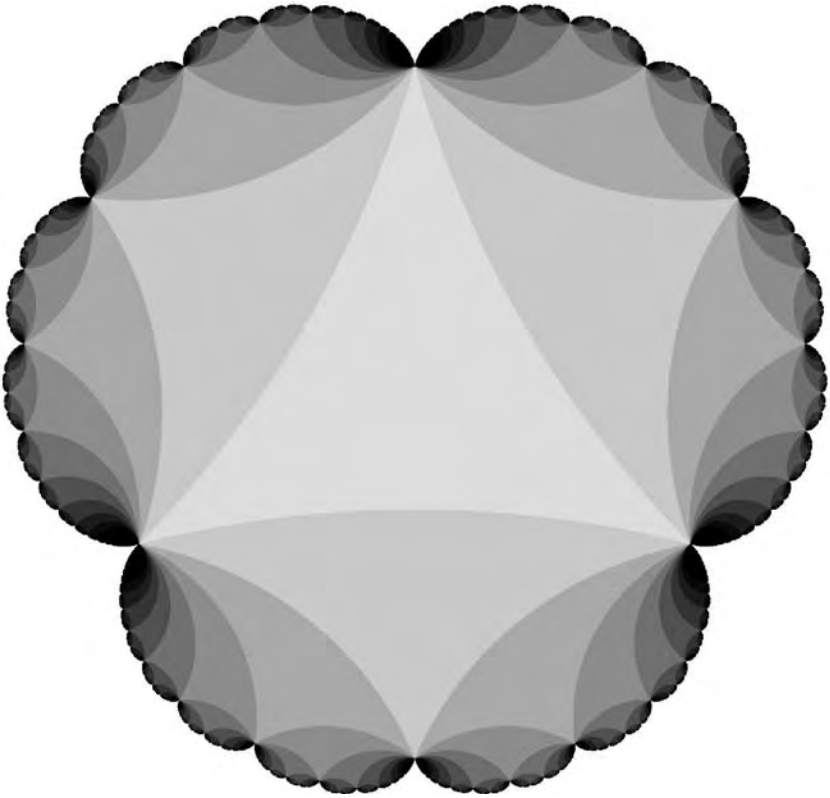


Fig. 5.5. Bers slice based on the hexagonal torus, inside the once-punctured torus space.

Bers slices are not the only $(3g+b-3)$-dimensional slices of quasifuchsian space. There are oodles of others. One can take a slice based on any boundary group of quasifuchsian space, except a doubly degenerate group. The bottom component S_{bot} might be a cusp, especially a maximal cusp (composed of triply punctured spheres). Such a slice is called a *Maskit slice*. One could as well take S_{bot} to be a singly degenerate end (in view of Sullivan's Theorem). Or one could require that $S_{\text{bot}} = S^{\text{top}}$ but that J match each $\gamma \in \pi_1(S^{\text{top}})$ with $\alpha(\gamma) \in \pi_1(S_{\text{bot}})$ where α is a fixed automorphism taken on all surfaces in the Teichmüller space. For example one could start with a reflection in the top surface of a fuchsian group arranged so that the positive imaginary axis is fixed, followed by the reflection in the real axis. Such submanifolds are called *Earle slices*.

5.8 The quasifuchsian space boundary

In this section we consider the closure $\overline{\mathfrak{T}(G)}$ of a quasifuchsian space $\mathfrak{T}(G)$. This is contained in $\mathfrak{R}_{\text{disc}}(G)$. As usual, denote its boundary by $\partial\mathfrak{T}(G)$.

The groups on $\partial\mathfrak{T}(G)$, referred to as boundary groups, or in this quasifuchsian context as *B-groups*, are classified as cusps (geometrically finite groups), singly degenerate or doubly degenerate groups, or partially degenerate groups.

We can approach a boundary cusp from $\mathfrak{T}(G)$ most directly by a process of pinching, as in Section 5.3 and Exercise 5-3.

In contrast, a singly or doubly degenerate group is a group H isomorphic to G such that either $\Omega(H)$ is connected or $\Omega(H) = \varnothing$. By Bonahon's criteria (page 253), $\mathcal{M}(H)$ is homeomorphic to $R \times (0, 1]$ in the singly degenerate case, or to $R \times (0, 1)$ in the doubly degenerate case. Either one component or the whole boundary has become degenerated. By the Density Theorem, all manifolds with this topological structure lie on $\partial\mathfrak{T}(G)$. By Sullivan's Theorem, a doubly generate group has no quasiconformal deformations (other than Möbius conjugations).

Doubly degenerate groups were discovered by Jørgensen on the boundary of once-punctured torus quasifuchsian space. His doubly degenerate groups H are periodic: there is a Möbius transformation $T \notin H$ satisfying $THT^{-1} = H$ such that the manifold $\mathcal{M}(H^*)$ corresponding to the augmented group $H^* = \langle H, T \rangle$ has finite volume. The the limit set of a doubly degenerate group is all $\mathbb{S}^2$. Yet any fiber $R \times \{s\}$, $0 < s < 1$, lifts to a planar object P in $\mathbb{H}^3$ on which H acts as a surface group. Even so, ∂P is dense in $\mathbb{S}^2$. For more details see the Double Limit Theorem (page 314) and the discussion following it. A doubly degenerate group might well contain new parabolics.

The cases of singly and partially degenerate groups will be discussed on page 270, again in the context of the Bers boundary.

The boundary has a lot of self-bumping as described in Theorem 5.3.2 because there are a lot of independent essential cylinders. The self-bumping occurs at most cusps. However there can be no self-bumping at maximal cusps (see Exercise 6-9). Nor (per Bromberg and Holt) is there any bumping at singly or doubly degenerate groups, with or without parabolics — geometric limits agree with algebraic at such points.

Collapsing mappings

The story here is based on [Minsky 1994a; 1994b]. We start with a fuchsian group G group acting in UHP and LHP and an isomorphism $\theta : G \to H$ to a discrete group H.

Consider the following test case. Take a simple closed geodesic $\gamma \in R' = \mathrm{UHP}/G$, and the set of lifts $\{\gamma_n^*\} \in$UHP. Suppose H is the cusp arising from R' by pinching γ. There is a continuous equivariant map $h : \mathbb{S}^1 \equiv \mathbb{R} \cup \infty \to \Lambda(H)$ which sends the endpoints of each γ_n^* to a single point but is otherwise one-one.

Now let Λ_+ and Λ_- be the ending laminations for $\mathcal{M}(H)$ as represented in UHP and LHP, respectively. We allow that one but not both laminations are empty if H is quasifuchsian. If both are nonempty, the leaves of Λ_- have no endpoints in common with the leaves of Λ_+ (Double Limit Theorem, page 314).

Theorem 5.8.1 [Minsky 1994a]. *Assume that G is a closed surface group and H has bounded geometry. There exists a continuous map $h : \mathbb{S}^2 \to \mathbb{S}^2$ which is equivariant*:

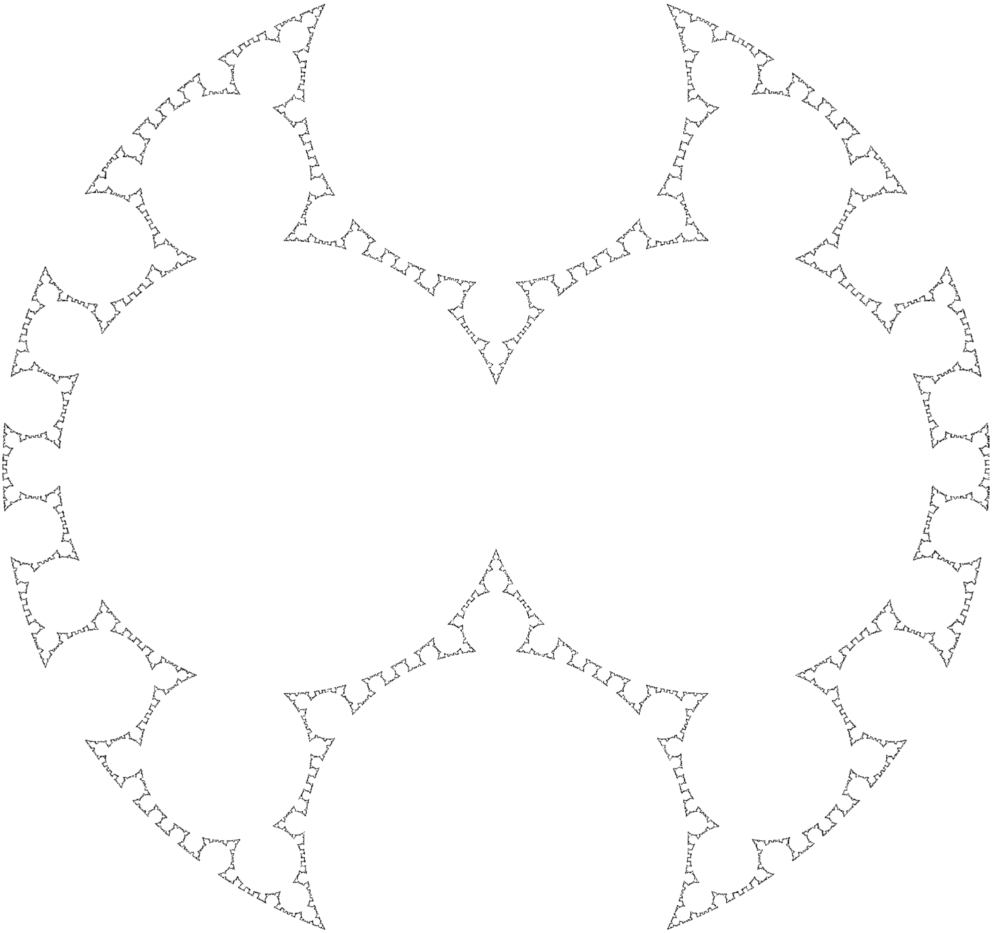


Fig. 5.6. The limit set of a 2-generator quasifuchsian group with commutator elliptic of order 3 and no parabolics.

$h \circ g(z) = \theta(G) \circ h(z)$ *for all* $g \in G$ *and* $z \in \mathbb{S}^2$. *It collapses the leaves to points* ($h(u) = h(v)$ *with* $u \neq v$) *if and only if* u *and* v *lie on the closure of the same leaf of* Λ_+, *or of* Λ_-. *If say* $\Lambda_- = \varnothing$ *then* h *is a homeomorphism of* LHP.

Thus if a component P of UHP$\setminus\Lambda_+$ is an ideal polygon, then $h(P)$ is a single point. See also Exercise 5-9.

As the limit set in the singly degenerate case is the continuous image of $\mathbb{S}^1$, we conclude at once from [Pommerenke 1992, Theorem 2.1]:

Corollary 5.8.2 [Minsky 1994a]. *Under the same hypotheses, if* H *is singly degenerate,* $\Lambda(H)$ *is locally connected.*

Minsky conjectured that Theorem 5.8.1 holds without the requirement of bounded geometry. McMullen [2001] proved that indeed, this is true for once-punctured tori. His argument makes heavy use of Minsky's model manifolds.

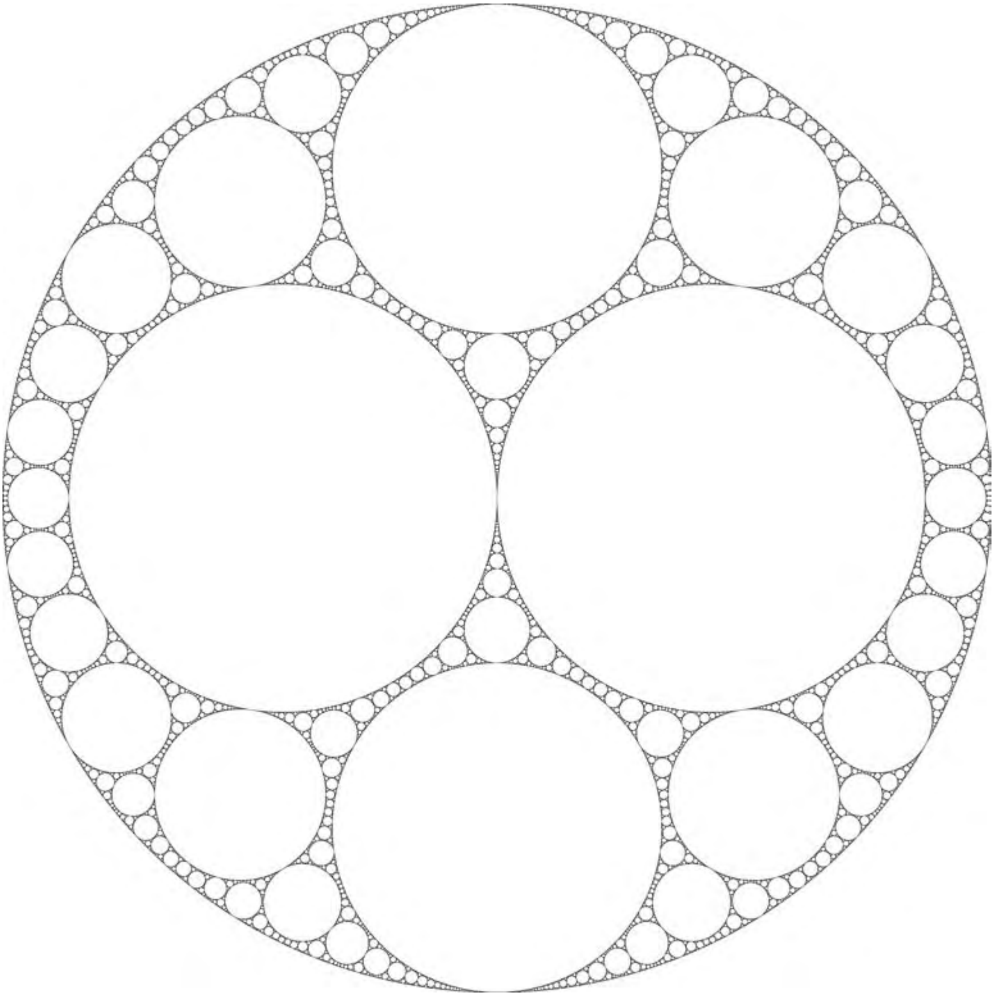


Fig. 5.7. A double cusp on the boundary of the quasifuchsian space of Figure 5.6. The stabilizer of each disk is a $(3, \infty, \infty)$ triangle group. Compare with Figure 5.9.

It is unknown whether all limit sets $\neq \mathbb{S}^2$ of kleinian groups are locally connected, although it is known in various special cases. This appears to be a very hard problem.

Predating Minsky's work is an amazing theorem of Cannon and Thurston showing how Peano curves (curves dense in $\mathbb{S}^2$) arise in the context of closed hyperbolic manifolds which fiber over the circle.

First consider the case that both ends of $\mathcal{M}(H)$ are infinite. We now have to assume that the ending laminations are those associated with the attracting and repelling fixed points of a pseudo-Anosov automorphism of $\mathbb{H}^2/G$ (see Exercise 5-6 and page 314). Take the closure of the lifts $\Lambda_-^* \subset \overline{\mathrm{LHP}}$ and $\Lambda_+^* \in \overline{\mathrm{UHP}}$. Construct a new space $\mathcal{S}$ from $\mathbb{S}^2$ as follows.

The points of $\mathcal{S}$ are (i) the components of $\overline{\mathrm{UHP}} \setminus \Lambda_+^*$ and the components of $\overline{\mathrm{LHP}} \setminus \Lambda_-^*$, (ii) the leaves of Λ_+^* and the leaves of Λ_-^*, and (iii) the points on $\mathbb{S}^2$ that are not in (i) or (ii). In forming the identification space $\mathcal{S}$ we have used the fact that the subsets of $\mathbb{S}^2$ we have used form a "cellular decomposition" because of the special nature of the ending laminations. The celebrated cellular decomposition theorem of R.L. Moore says that $\mathcal{S}$ is *homeomorphic* to $\mathbb{S}^2$ itself! We thus get a *continuous map* $f : \mathbb{S}^2 \to \mathbb{S}^2$ by projecting from $\mathbb{S}^2$ to the identification space $\mathcal{S}$ and then onto $\mathbb{S}^2$. The map f conjugates G to H, that is $f \circ g(x) = \theta(g) \circ f(x)$ for some isomorphism θ. Therefore $f(\mathbb{S}^1)$ is an H-invariant Peano curve. We think of this curve as resulting from the collapse of the two laminations.

On other hand suppose the bottom end of $\mathcal{M}(H)$ is homeomorphic to $\mathbb{H}^2/G$ while Λ_+, the ending lamination of the top, corresponds to the attracting or the repelling fixed point for a pseudo-Anosov map. Form the space $\mathcal{S}$ using (i) and (ii) in $\overline{\text{UHP}}$ but replace (iii) by (iii') points in $\overline{\text{LHP}}$ which are not in (i) or (ii). Again the identification space $\mathcal{S}$ is homeomorphic to $\mathbb{S}^2$ resulting in a continuous map $f : \mathbb{S}^2 \to \mathbb{S}^2$ that conjugates G to H and is a homeomorphism on LHP. In this case $f(\mathbb{S}^1)$ is the limit set of H; the limit set results from collapsing the ending lamination.

A less specific collapsing theorem that parallels the geometrically finite case was proved by Erica Klarreich (her original statement required tameness):

Theorem 5.8.3 [Klarreich 1999]. *Assume G is freely indecomposable and $\mathcal{M}(G)$ is geometrically finite. Suppose there is a homeomorphism $f : \mathbb{H}^3/G \to \mathbb{H}^3/H$, where $\mathbb{H}^3/H$ has bounded geometry. Then there is a homotopic homeomorphism $h \sim f$ whose lift to $\mathbb{H}^3$ extends to be a continuous, surjective, equivariant map $\mathbb{S}^2 \to \mathbb{S}^2$.*

For simply or doubly degenerate groups without new parabolics, bounded geometry is determined by a property of the ending laminations [Minsky 2001].

The Bers boundary

The Bers slice $\mathcal{B}(R) \equiv \text{Teich}(R)$ has a boundary $\partial\mathcal{B}(R) \equiv \partial\mathfrak{T}(G) \subset \mathfrak{R}(G)$, called the *Bers boundary* or *analytic boundary*. Unlike the boundary of the full quasifuchsian space, $\mathcal{B}(R) \cup \partial\mathcal{B}(R)$ is compact. This is because the family of normalized conformal maps of LHP that conjugate G to another group H is compact. Therefore for every boundary group H, exactly one component Ω_{bot} of $\Omega(G)$ is invariant under the full group H. The theory of the boundary was first worked out in the pioneering papers [Bers 1970a] and [Maskit 1970].

In the case that G represents a once-punctured torus so that $\mathcal{B}(R)$ has complex dimension one, it is known [Minsky 1999] that $\partial\mathcal{B}(R)$ is a Jordan curve (in a planar embedding), and hence locally connected. Also see [McMullen 1998].

If a boundary group $H = \theta(G)$ is geometrically finite, that is if H is a cusp, $\mathcal{M}(H)$ has the following structure. There are a finite number of mutually disjoint simple closed geodesics $\{\alpha_i\}$ on $S_{\text{bot}} = \partial_{\text{bot}}\mathcal{M}(H) = \Omega/H$, which determine parabolic transformations in H. These divide S_{bot} into one or more components $\{S_i\}$. Each S_i is parallel in $\mathcal{M}(H)$ to a component S_i' of $\partial^{\text{top}}\mathcal{M}(H)$ which is a finitely punctured closed surface homeomorphic to S_i. The stabilizing subgroup of a component Ω_i' over S_i' is

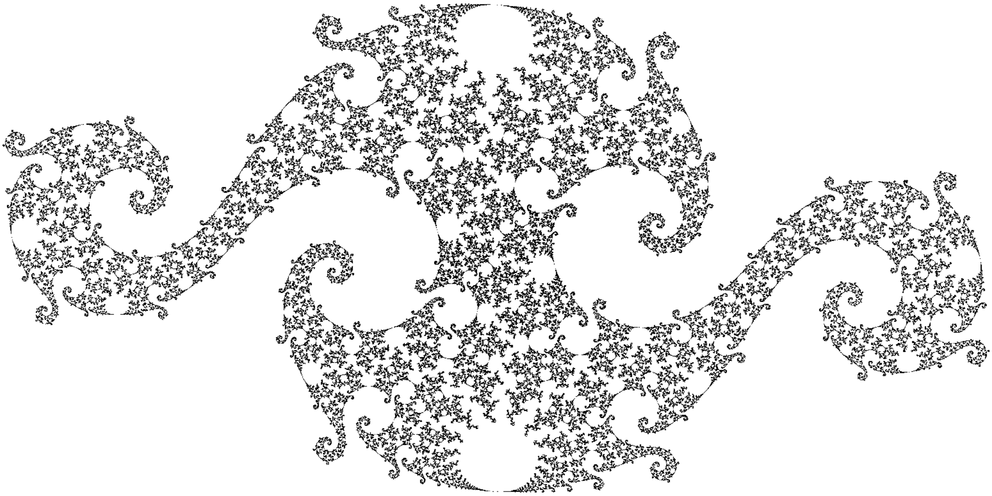


Fig. 5.8. This beautifully crafted limit set of Hausdorff dimension two belongs to a singly degenerate once-punctured torus group on the boundary of a Bers slice. The complement is connected (can you find your way out of the maze?) The "holes" in the picture contain horodisks at the parabolic fixed points and are an artifact of the algorithm: extremely long words in the generators would be required to fill them in.

quasifuchsian. The configuration $\bigcup S_i'$ results from pinching the top surface in the free homotopy classes of the geodesics in S_{bot}.

In Ω^{top}, as a loxodromic element of G becomes parabolic in H, its two fixed points coalesce into one, pinching Ω^{top} into two simply connected pieces. This happens simultaneously in the conjugacy class of g so that Ω^{top} becomes pinched off into countably many simply connected regions lying over the punctured surfaces $\{S_i'\}$. If we focus on the geodesics in the original manifold $\mathcal{M}(G)$ that are in the free homotopy classes of $\{\alpha_i\}$ then as we approach $\mathcal{M}(H)$ these and only these "exit" $\partial^{\mathrm{top}}\mathcal{M}(H)$ in the sense that they are "becoming" parabolic fixed points.

For more details see [Maskit 1970; 1988; Marden 1974a; 1977] and Exercise 5-3.

Recall that a maximal cusp on the Bers boundary is a cusp for which $\partial\mathcal{M}^{\mathrm{top}}$ is a union of triply punctured spheres. It corresponds to a set of pinching loops which form a pants decomposition of S_{bot} — each complementary component is a 3-holed sphere.

One or all of the top surfaces S_i' can themselves be degenerated so long as they are not triply punctured spheres. When not all are degenerated, the group is referred to as a *partially degenerated group*. Even if all become degenerated, one may be left with parabolic transformations which on Ω_{bot} act as hyperbolic transformations. If S^{top} is entirely degenerated with or without any such pinchings, the resulting boundary group is *singly degenerate*. Singly degenerate groups are characterized by the property that $\Omega(G)$ is connected and simply connected. Such groups are constructed in Section 6.1.1.

Suppose G has no parabolics. There exist boundary groups H without any parabolics either, in fact there are lots of boundary groups which have no parabolics. For as pointed out by Bers [1970a], since G has a countable number of elements, the set of

$$H = \theta(G) \in \partial\mathcal{B}(R)$$

for which $\theta(g)$ is parabolic for some g has positive codimension in $\partial\mathcal{B}(G)$. Most boundary groups H are without parabolics, and therefore singly degenerate.

By Bonahon's criteria (page 253), the interior of the manifold coming from a boundary group is, like all other manifolds of the deformation space, homeomorphic to $S_{\text{bot}} \times (0, 1)$. The long standing question as to whether, conversely, every kleinian group with the structure of a singly degenerate group is a boundary group of some Bers slice was first answered affirmatively as a special case of the density conjecture:

Bromberg's Theorem. *Suppose Γ has no parabolics, is isomorphic to the fundamental group of a closed surface, and $\mathcal{M}(\Gamma)$ has exactly one geometrically finite end. Then Γ is a boundary point of the Bers slice determined by its geometrically finite end. Moreover, Γ is the algebraic limit of a sequence of quasifuchsian groups lying in this Bers slice.*

This was extended to all quasifuchsian spaces by Brock and Bromberg, and finally incorporated into the framework of the Density Theorem (p. 260).

McMullen [1991] proved the strongest form of a longstanding conjecture of Bers and introduced ideas that have been used for the more general Theorem 5.3.1.

Theorem 5.8.4 (Maximal cusps are dense, I). *In the topology of algebraic convergence*, *maximal cusps are dense on the boundary of any Bers slice $\mathcal{B}(R)$ in quasifuchsian space*, *where R is a Riemann surface of finite hyperbolic area*.

It was established in [Kerckhoff and Thurston 1990] that when the genus of R exceeds one, the natural (biholomorphic) map from one Bers slice to another does *not* have a continuous extension to a map between the corresponding Bers boundaries. For the once punctured torus case, on the other hand, the map does extend continuously to the boundaries.

5.9 Geometric limits at boundary points

We ask, if a sequence converges algebraically to a boundary point H of a quasifuchsian deformation space, what are its possible geometric limits?

We know from Theorem 4.6.2 that if $\Omega(H) = \varnothing$, then any geometric limit coincides with the algebraic limit. This is also true when $\Omega(H) \neq \varnothing$, provided the algebraic limit has no new parabolics.

Suppose $X \in \partial\mathcal{T}(G)$ is a maximal cusp. It is also true that for any geometric limit X^* at X, the convex core of $\mathcal{M}(X)$ is embedded in $\mathcal{M}(X^*)$; see [Anderson et al. 1996, Prop. 32].

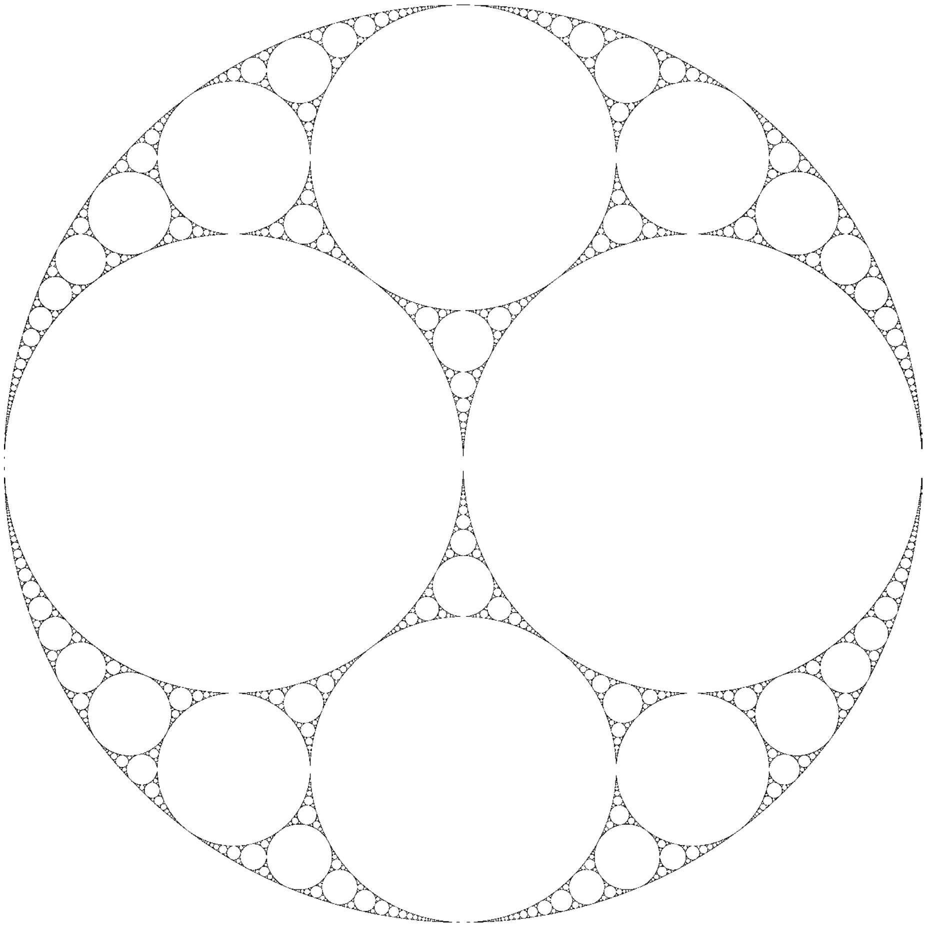


Fig. 5.9. The limit set of a maximal cusp (double cusp) on the boundary of the once-punctured torus quasifuchsian space. The limit set can alternatively be constructed by the iterative process of inscribing the maximal circle in each triangular interstice. It is called the apollonian gasket, because it was first constructed in the third century B.C. by the Greek mathematician Apollonius [Mumford et al. 2002]. The tangencies are at parabolic fixed points.

The Jørgensen picture of the once punctured torus case

It is illuminating to consider the case of the once-punctured torus quasifuchsian space, where the possibilities were enumerated by Jørgensen in unpublished work. There, an end of the manifold corresponding to a boundary group is either degenerated, or it is a thrice punctured sphere. Only in the latter case can a geometric limit be strictly larger than the algebraic.

It is easiest to understand Jørgensen's description if we start with a cusp group $H = \theta(G)$ such that $\partial\mathcal{M}(H)$ consists of two thrice punctured spheres; we will refer to this as a double cusp group. Of the three punctures on each component, one corresponds

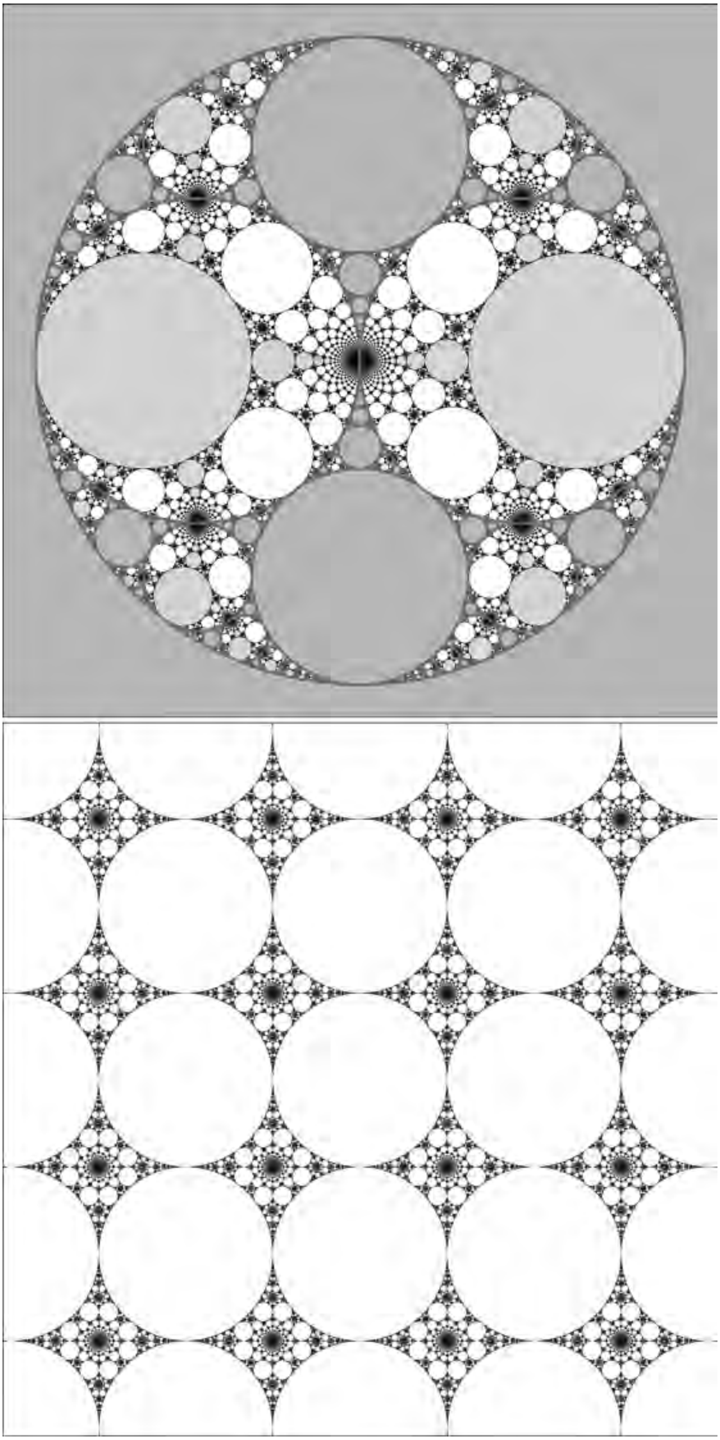


Fig. 5.10. A geometric limit at the group of Figure 5.9. The shaded disks are paired by generators of the rank two conjugacy classes. The alternate view at the bottom results from placing a rank two cusp at ∞.

to the puncture on the original two once-punctured boundary tori; the original pair of punctures and its pairing tube serves as a kind of backbone for $\mathcal{M}(H)$.

Corresponding to each triply punctured sphere boundary component S_1, S_2, there is a generator pair A_i, B_i of H with the following properties. Each commutator $[A_i, B_i]$ generates a rank one parabolic subgroup whose conjugacy class forms the backbone. The elements A_i are parabolic, corresponding to the pair of new parabolics on S_i, $i = 1, 2$, and generate nonconjugate subgroups. The elements B_i are loxodromic. Such generator pairs can be found by inserting the solid pairing tube associated with the two new punctures on each component so forming again a once-punctured torus. The generator pairs so associated with the two boundary components are related by Nielsen transformations as in Exercise 5-10. A double cusp group is uniquely determined by nonconjugate generator pairs $\langle A_1, B_1\rangle$, $\langle A_2, B_2\rangle$.

According to Jørgensen, all geometric limits $\Gamma \supset H$ correspond to manifolds $\mathcal{M}(\Gamma)$ which are fibered, fibered either over $\mathbb{R}$, or over the half open or closed intervals $[0, +\infty)$, $(-\infty, 1]$, $[0, 1]$ depending on which end or ends are thrice punctured spheres. In explaining the picture let's concentrate on the doubly infinite case.

In this case there is a countably infinite number of "singular" fibers in $\mathcal{M}(\Gamma)$ representing thrice punctured spheres, while all the other fibers are once-punctured tori. All the fibers of $\mathcal{M}(\Gamma)$ connect to the backbone.

The singular fibers divide $\mathcal{M}(\Gamma)$ into chunks fibered by nonsingular fibers. A chunk corresponds to a double cusp subgroup $\langle A, B\rangle \subset \Gamma$ that represents a once-punctured torus and has the following property. The result of pinching A gives the singular fiber at one end of the chunk, and pinching B gives the singular fiber at the other end. The groups $\{\langle A, B\rangle\}$ corresponding to different chunks are not conjugate in Γ. In particular, Γ is not finitely generated.

Each singular fiber S has two new punctures. The are joined by a solid pairing tube τ_1 in the chunk on one side, and τ_2 in the chunk on the other. The two tubes join together to form a rank two solid cusp torus in the union of the two adjacent chunks.

For an explicit construction of groups of this type see Exercise 5-12. The analogue for general quasifuchsian groups is presented in [Thurston 1986c, Theorem 7.2].

It is relevant to cite [Anderson et al. 1996, Prop. 3.2] wherein the following is observed for a geometrically finite G: Assume $X \in \mathfrak{T}(G)$ is the algebraic limit of a sequence in $\mathfrak{T}(G)$. Suppose X is a maximal cusp. Then for any geometric limit $X^* \supset X$ of a subsequence, the convex core of $\mathcal{M}(X)$ is embedded in $\mathcal{M}(X^*)$.

Geometric limits at the Bers boundary

We have already indicated in the context of the Jørgensen picture one class of geometric limits. More generally, suppose the group H represents a cusp on a $\partial\mathcal{B}(G)$. The Jørgensen picture suggests that if H is approached "radially" from within $\mathcal{B}(G)$, the geometric and algebraic limits coincide. To get a larger geometric limit, it is necessary to approach the cusp "tangentially" or even "ultratangentially". By analogy, in the modular group contrast the radially approach $z_n \to \zeta$ to the fixed point of a parabolic T with the tangential approach $z_n = T^n(z) \to \zeta$.

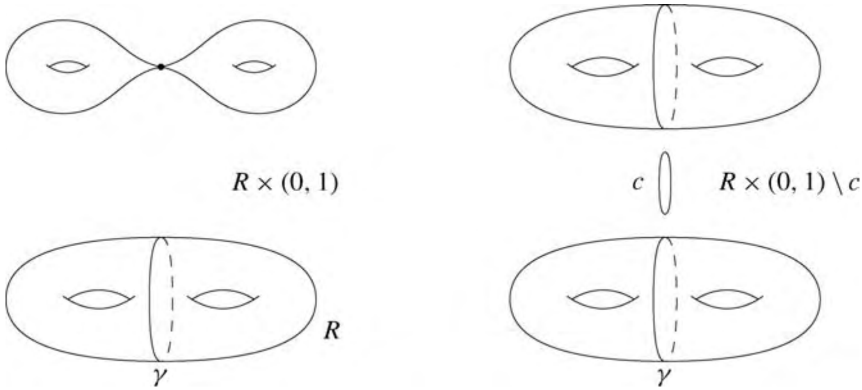


Fig. 5.11. The algebraic (left) and geometric (right) limits at a cusp on the boundary of the quasifuchsian space of a closed, genus-2 surface obtained as the limit of the iteration of a point O in a Bers slice by powers of the Dehn twist about γ. The top surface of the algebraic limit is the result of pinching $\gamma \subset R^{\text{top}}$ of O while the top surface of the geometric is conformally equivalent to R^{top}. See Exercises 4-20, 5-6, 5-11.

Brock [2001a] has given a complete description of the geometric limits resulting from the iteration $\{\tau^n(O)\}$ of a point $O \in \mathcal{B}(G)$ by a reducible automorphism τ. This generalizes the case that τ is a Dehn twist (Exercise 5-5), which was examined in [Marden 1980; Kerckhoff and Thurston 1990] and is presented in Exercise 5-11.

A *reducible* automorphism (see Exercise 5-6) is an automorphism $\tau : R \to R$ that fixes a set of free homotopy classes represented by mutually disjoint, nontrivial, simple closed curves $\{\gamma_i\}$ on the surface R, none of which can be homotoped to a puncture. We may assume the curves themselves are fixed: The complementary components are neither simply nor doubly connected. Some power τ^s fixes each free homotopy class and consequently each complementary component $R \setminus \bigcup \gamma_i$. We may assume that in each complementary component τ is either (homotopic to) the identity, a map of finite order, or is pseudo-Anosov — that is no free homotopy classes of simple curves are fixed, except those of the boundary components. Therefore we may assume that τ not only preserves the components and the boundary curves, but in each component it is either homotopic to the identity or is pseudo-Anosov. An automorphism τ is called a *reducible pseudo-Anosov* if it is pseudo-Anosov in at least one complementary component. If τ acts as the identity on both sides of a loop γ_j, then τ is a Dehn twist (Exercise 5-5) about γ_j. The once-punctured torus is special in that there are no reducible pseudo-Anosovs.

Now let τ act on the Bers slice $\mathcal{B}(R)$. Choose a point $O \in \mathcal{B}(R)$ and denote its top surface by S and bottom, which is constant throughout the slice, by R. Let $\{R_{\text{pA}}\}$ denote the components of $R \setminus \bigcup \gamma_i$ on which τ is pseudo-Anosov and $\{R_{\text{id}}\}$ the remaining components, if any, on which τ acts as the identity. Denote the subsurfaces of S parallel to them by $\{S_{\text{pA}}\}$ and $\{S_{\text{id}}\}$. In the algebraic limit, the components of

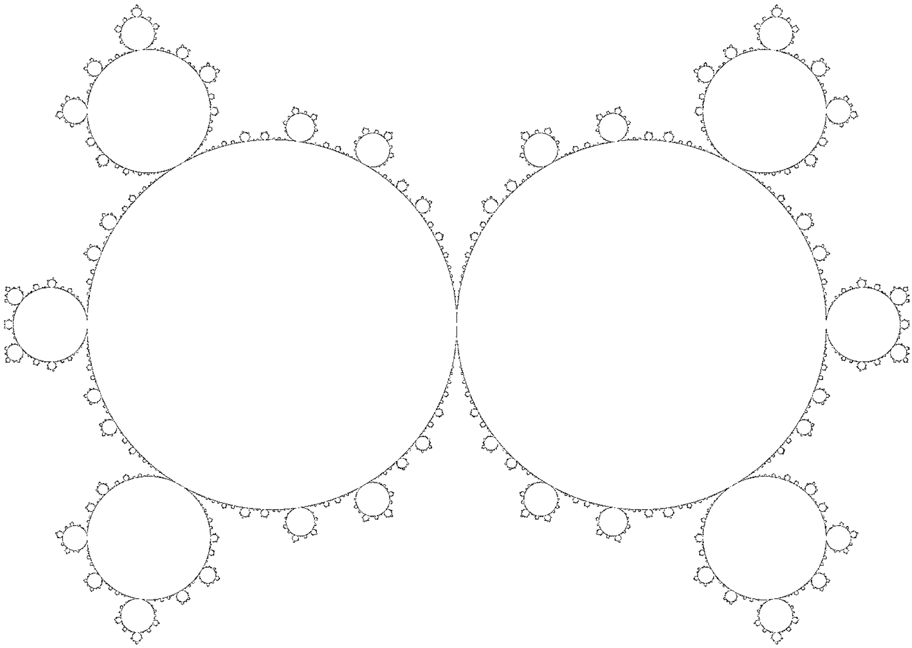


Fig. 5.12. The limit set of an algebraic limit corresponding to Figure 5.11. The cusp is chosen so that the pinched components of the manifold are covered by round disks.

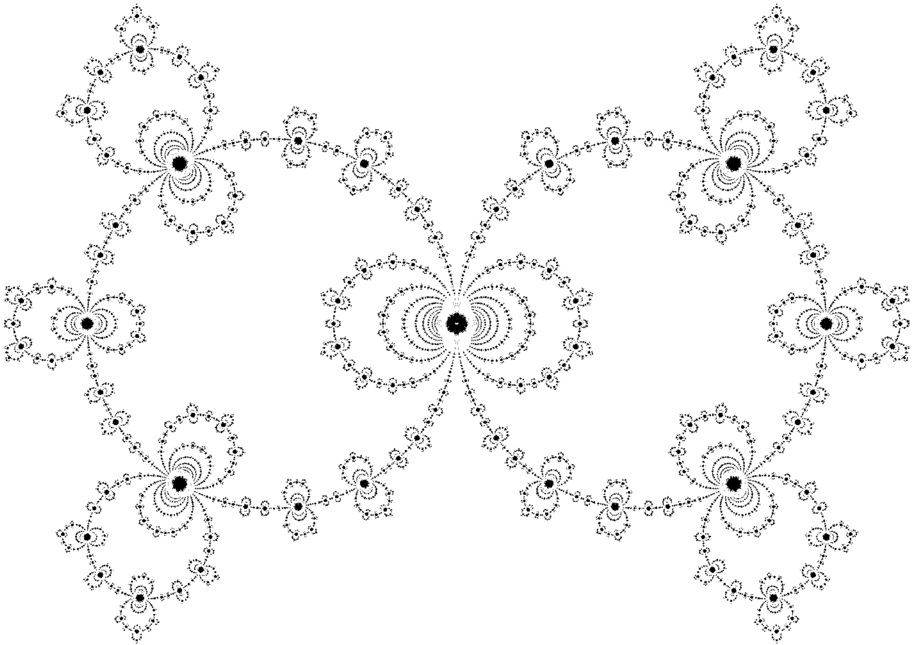


Fig. 5.13. The limit set of the geometric limit at the cusp of Figure 5.12. It can be constructed by reflecting the algebraic limit in the circles.

$\{S_{\text{pA}}\}$ are headed for degeneration and those of $\{S_{\text{id}}\}$ are due to have their boundary components pinched. In fact all of the loops $\{\gamma_j\}$ will be pinched.

For simplicity of description, assume there is one component of each set, namely S_{pA} and S_{id}. Also assume that their common boundary is formed by a single simple loop γ. The parallel subsurfaces of R, namely R_{pA} and R_{id}, are bounded by a loop parallel to γ. Suppose the basepoint O corresponds to the quasifuchsian group G.

Theorem 5.9.1 [Brock 2001a]. *The sequence of iterates $\{\tau^n(O)\}$ described above converges algebraically to a group $\varphi : G \to H \in \partial\mathcal{B}(G)$ and geometrically to a group H^* properly containing H. These have the following properties*:

(i) *The isomorphism φ sends the cyclic loxodromic subgroup of G corresponding to each lift of γ to a rank one parabolic subgroup of H.*
(ii) *φ is associated with a conformal map Φ between the bottom surfaces R of $\partial\mathcal{M}(G)$ and R of $\partial\mathcal{M}(H)$. and a homeomorphism $\Phi : S_{\text{id}} \to \partial\mathcal{M}(H) \setminus R$; γ corresponds to the puncture on $\Phi(S_{\text{id}})$.*
(iii) *The H-stabilizer of each lift of R_{id} to $\mathbb{H}^3$ is a quasifuchsian subgroup of H; the H-stabilizer of each lift of $\{R_{\text{pA}}\}$ is a singly degenerate subgroup of H.*
(iv) *$\partial\mathcal{M}(H^*)$ has two boundary components, one conformally equivalent to R the other conformally equivalent to S.*
(v) *The H^*-stabilizer of each lift of R_{id} and S_{id} is a quasifuchsian subgroup of H^*. The H^*-stabilizer of each lift of R_{pA} and of S_{pA} is a singly degenerate subgroup of H^*.*
(vi) *The interior $\mathbb{H}^3/H^*$ is homeomorphic to $R \times (0, 1) \setminus [R_{\text{pA}} \times \{1/2\}]$. Thus $\mathcal{M}(H^*)$ has two degenerate ends, corresponding to the two sides of $R_{\text{pA}} \times \{1/2\}$.*

There is an analogous description for the general case. The only difference occurs when both sides of a loop γ_j belong to S_{id}, that is either γ_j is part of the common boundary of two components, or is a pinching loop of a single component. In either case γ_j determines a rank two parabolic subgroup of the geometric limit H^*, as described in Exercise 5-11.

Brock showed that by applying the techniques of the Skinning Lemma (Section 6.2), groups with the properties of Theorem 5.9.1 can be directly constructed as follows: Find groups $H_1, H_2 \in \partial\mathfrak{T}(G)$ with $\partial\mathcal{M}(H_1) = R \cup S_{\text{id}}$, $\partial\mathcal{M}(H_2) = S \cup R_{\text{id}}$, that is degenerate S_{pA}, R_{pA}, respectively. Then use the skinning lemma to identify R_{id} on the bottom of $\partial\mathcal{M}(H_2)$ with S_{id} on the top of $\partial\mathcal{M}(H_1)$. In the case that τ is a Dehn twist, this is explained in Exercises 5-7 and 5-11. See Figures 5.14, 5.15, 5.16.

The totality of geometric limits at the quasifuchsian space boundary

Let G denote a fuchsian group representing a closed surface of genus ≥ 2. Teruhiko Soma recently presented a complete description of the *topological* possibilities for geometric limits at boundary points of the full quasiconformal deformation space $\mathfrak{T}(G)$. The totality of possible geometric limits contains an amazing richness of structure, yet each limit is organized in slices parallel to the top and bottom. While they all by

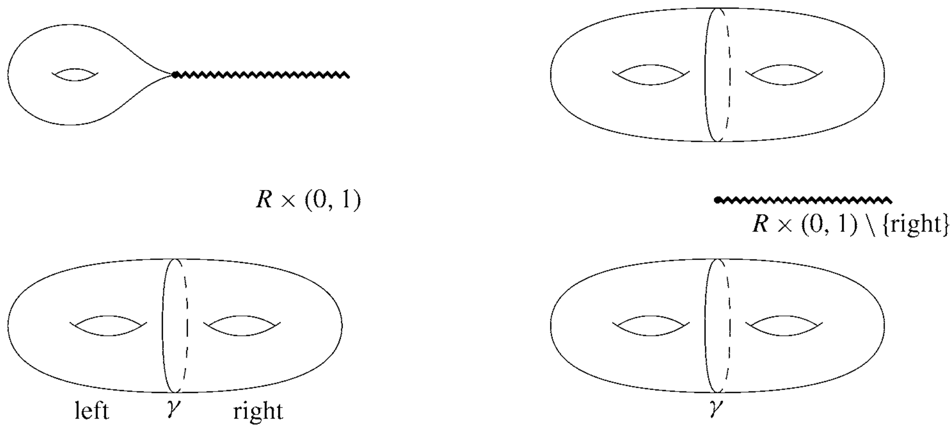


Fig. 5.14. The algebraic and geometric limit at a boundary point of a genus-2 Bers slice. In the algebraic limit, half the top surface has degenerated leaving a once-punctured torus. The geometric limit is homeomorphic to the result of removing a once-punctured torus from $R \times (0, 1)$. It is the result of iterating the right side by a pseudo-Anosov that fixes the left side pointwise, as in Theorem 5.9.1.

necessity contain new parabolics, our prior examples show that the new parabolics can appear both in rank one and in rank two parabolic subgroups. It appears that the Minsky bilipschitz models of hyperbolic manifolds lead to a geometric description of the geometric limits.

Now $\mathcal{M}(G) \cong S \times I$, $I = [0, 1]$. We will refer to the slices $S_y = q^{-1}(y) = S \times \{y\}$ corresponding to the projection $q : S \times I \to I$ and put the hyperbolic metric on each one.

Theorem 5.9.2 [Soma 2003]. *At a boundary point H of $\mathfrak{T}(G)$ let H^* be a geometric limit which is strictly larger than H. There exists a closed set $\mathcal{X} \subset S \times I$ containing the top and bottom* $S \times \{0\}$, $S \times \{1\}$, *with the property that its complement $S \times I \setminus \mathcal{X}$ is homeomorphic to $\mathbb{H}^3/H^*$ and contains $S \times \{\frac{1}{2}\}$.*

Each slice $X_y = \mathcal{X} \cap S_y$, $y \in \mathcal{Y} = q(\mathcal{X})$, is the disjoint union of a compact subsurface with geodesic boundary components and simple geodesics (*either set may be empty*).

The set $\mathcal{X}$ has additional properties spelled out in [Soma 2003]. The set $\mathcal{Y} \subset I = q(\mathcal{X})$ is not discrete in general and may contain intervals. When $\mathcal{Y}$ is totally disconnected (each component is a point) then each connected component of $\mathcal{X}$ is either a subsurface or a simple closed geodesic in some S_y. Yet there still may be accumulation points — and accumulation points of accumulation points!

To study the accumulation to S_y, define for $y \in \mathcal{Y}$, and $y < 1$, $y > 0$ respectively:

$$\Lambda_y^+ = S_y \cap \left(\mathcal{X} \cap \overline{S \times (y, 1]}\right), \quad \Lambda_y^- = S_y \cap \left(\mathcal{X} \cap \overline{S \times [0, y)}\right).$$

For example $\Lambda_y^\pm$ may be the geodesic lamination arising as the accumulation of a sequence of simple geodesics in other levels $\{y\}$. If the geodesics have positive intersection numbers with each other (upon projection to S_0), one cannot slide past

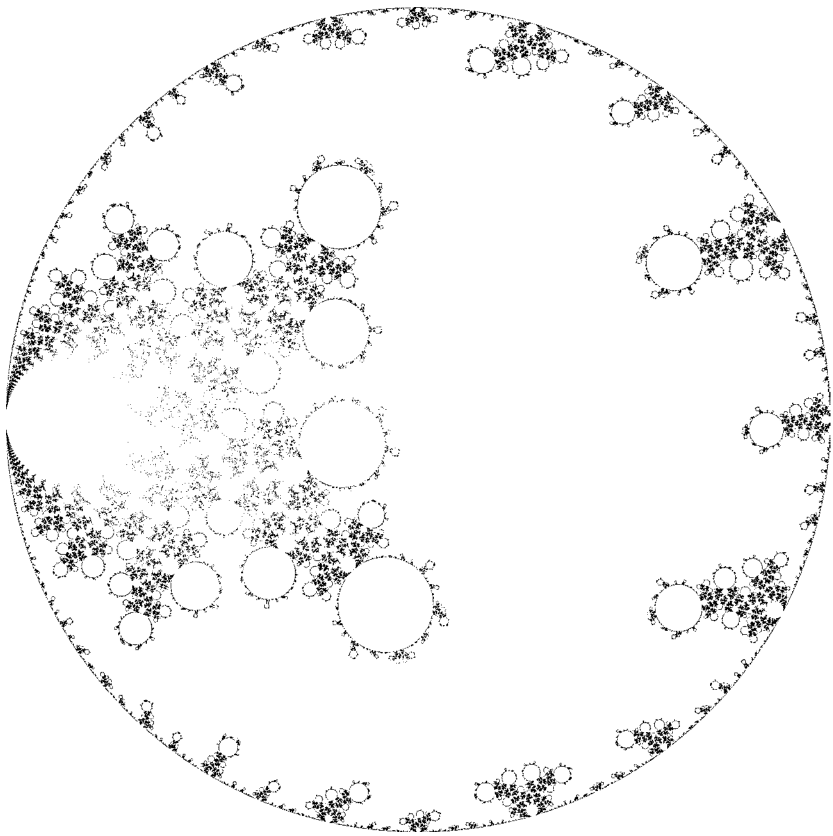


Fig. 5.15. The limit set of an algebraic limit corresponding to Figure 5.14. The boundary point is chosen so that the good half of the top is covered by a round disk.

another; in particular the sequence can be prevented from converging to the top or bottom S_0, S_1 (unlike the Jørgensen example). One also needs to consider certain closed subsets of S_y denoted by $\Delta(\Lambda_y^{\pm}) \supset \Lambda_y^{\pm}$. For each $\epsilon = \pm$, the relative boundary $\partial\Delta(\Lambda_y^{\epsilon})$ is to be the disjoint union of simple closed geodesics $\partial F \cup \ell_1 \cup \cdots \cup \ell_m$, where $F \subset S_y$ is a geodesic subsurface. The simple geodesics in $\Delta(\Lambda_y^{\pm})$ having a one sided open annular neighborhood not meeting $\Lambda_y^{\pm}$ correspond to parabolic elements of H^*.

Soma also shows how groups H^* described in the theorem can be constructed.

The Thurston boundary

The main references for this section are [Thurston 1986c; Fathi et al. 1979, §8; Bonahon 1988].

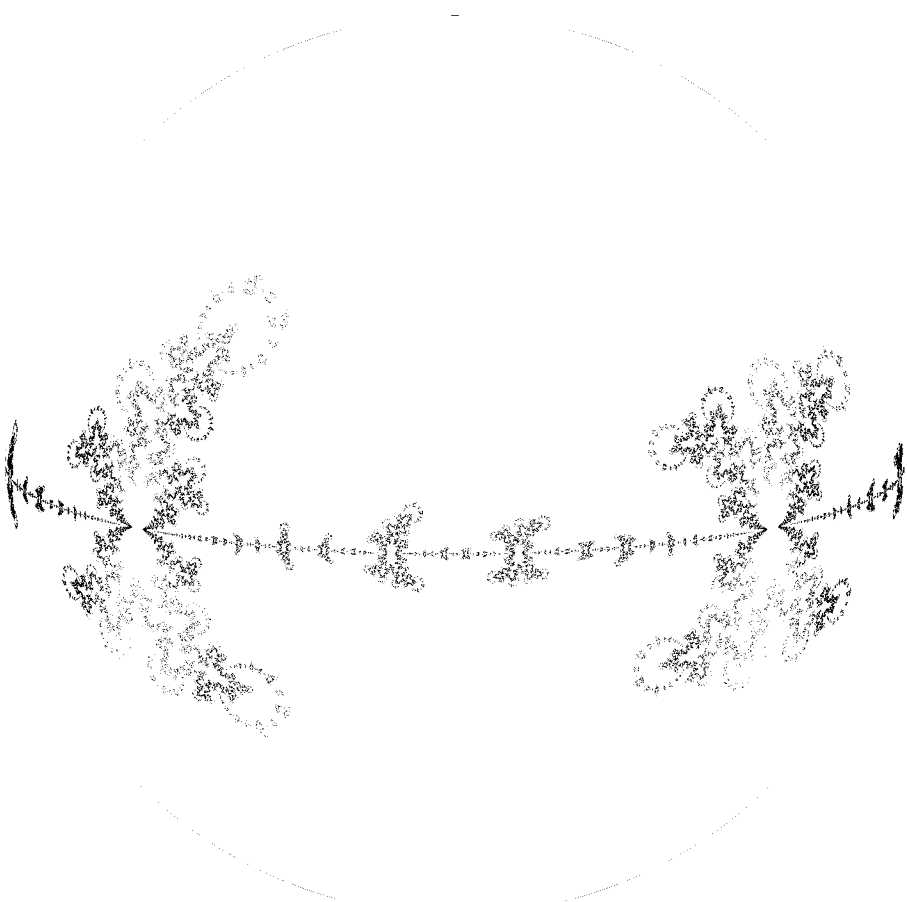


Fig. 5.16. The limit set of the geometric limit corresponding to Figure 5.14. It can be obtained by reflecting the algebraic limit set in the round circles.

Fix a hyperbolic (Riemann) surface R as our reference surface. Let γ be a simple loop on R (not homotopic to a point). Denote by $\ell_\rho(\gamma)$ the length of the geodesic freely homotopic to γ in the hyperbolic metric ρ of R. Thurston expresses the topology of Teich(R) as the minimum topology that for any fixed γ, $\ell_\rho(\gamma)$ is a continuous function of ρ. Following [Thurston 1988], consider the projectivized functional $L : \text{Teich}(R) \to \mathcal{PML}(R)$ that sends each $\gamma \subset R$ to $\ell_\rho(\gamma)$ *modulo* positive multiplicative constants. He asserts that the closure of its image is Teich$(R) \cup \mathcal{PML}(R)$, which is homeomorphic to the closed ball $B^{6g+2b-6}$. The boundary is called the *Thurston boundary*. We will explain in more detail.

Put a succession of new hyperbolic metrics $\{\rho_n\}$ — new complex structures — on R. We can more generally find the length $\ell_{\rho_n}(\nu)$ of the ρ_n-geodesic lamination determined by a given geodesic lamination $\nu \subset R$, see §3.9.1. For example, ρ_n may come from the hyperbolic metric on a pleated surface in some $\mathcal{M}(G) \in \mathfrak{T}(R)$.

A sequence of hyperbolic structures $\{\rho_n\} \subset \text{Teich}(R)$ is said to *converge* to $(\Lambda, \mu) \in \mathcal{PML}(R)$ if and only if there is a sequence of positive numbers $\{c_n \to \infty\}$ such that for all $\nu \in \mathcal{ML}(R)$ with $\iota(\nu, \mu) \neq 0$,

$$\lim_{n\to\infty} \frac{\ell_{\rho_n}(\nu)}{c_n} = \iota(\nu, \mu).$$

It is enough to take the measures ν to be supported on simple closed geodesics. We can then express the convergence criterion as either of

$$\lim_{n\to\infty} \frac{\ell_{\rho_n}(a)}{c_n} = \iota(a, \mu), \qquad \lim_{n\to\infty} \frac{\ell_{\rho_n}(a)}{\ell_{\rho_n}(b)} = \frac{\iota(a, \mu)}{\iota(b, \mu)},$$

for any simple loop a, or pair a, b of simple loops on R with $\iota(b, \mu) \neq 0$. Note that μ is determined only up to a multiplicative constant.

For example, suppose Λ is a simple loop α and μ is its unit atomic measure, then $\ell_{\rho_n}(b) \to \infty$ for all simple loops b with $\iota(b, \alpha) \neq 0$.

Thurston [1986c, Theorem 2.2] proves that $\{\rho_n\}$ converges to a lamination $(\Lambda, \mu) \in \mathcal{PML}(R)$ if and only if there is a sequence of measured laminations $\{(\Lambda_n, \mu_n)\}$ converging projectively to (Λ, μ) such that for all $\nu \in \mathcal{ML}(R)$ with $\iota(\nu, \mu) \neq 0$,

$$\lim \frac{\ell_{\rho_n}(\nu)}{\iota(\mu_n, \nu)} = 1, \quad \ell_\rho(\mu_n) \to \infty, \;\; \ell_{\rho_n}(\mu_n) < C < \infty, \tag{5.3}$$

for some constant C and all indices.

Therefore $\widehat{\mu}_n = \mu_n / \ell_\rho(\mu_n)$ in the projective class of μ_n is such that $\ell_{\rho_n}(\widehat{\mu}_n) \to 0$. We can choose μ_n supported on a simple loop γ_n. In this case $\ell_{\rho_n}(\gamma_n)/\rho(\gamma_n) \to 0$. For a specific example, see Section 6.1.1.

It is not necessarily possible to choose μ_n in the projective class of the limit μ to satisfy all conditions of Equation (5.3). As an example let ρ_n denote the new hyperbolic metric on R that comes from the n-th iterate τ^n of the Dehn twist about a simple geodesic γ: $\ell_{\rho_n}(\tau^n(b)) = \ell_\rho(b)$, for all simple geodesics b. In particular, $\ell_{\rho_n}(\gamma) = \ell_\rho(\gamma)$. For some $\{c_n\}$, $\{\tau^n(b)/c_n\}$ converges in $\mathcal{PML}(R)$ to $\Lambda = \gamma$ for every b with $\iota(b, \gamma) \neq 0$. Equation (5.3) is satisfied for $\mu_n = (n\ell_\rho(\gamma))\gamma$ except $\ell_{\rho_n}(\mu_n)$ is not bounded. However, $\ell_\rho(\tau^n(b)) \to \infty$ only when $\iota(b, \gamma) \neq 0$.

With the topology suggested above, $\mathfrak{T}(R) \cup \mathcal{PML}(R)$ forms a compact metric space and its (compact) boundary $\partial_{\text{th}}\mathfrak{T}(R) = \mathcal{PML}(R)$ is called the *Thurston boundary*.

A Bers boundary point corresponds to a single Thurston boundary point provided the space of projective measures with support on the ending lamination has dimension zero. Thus a maximal cusp on the Bers boundary determined by $3g-3$ pinching curves corresponds to a $(3g-4)$-dimensional subspace on the Thurston boundary. The result of pinching a single curve gives rise to a $(3g-4)$ dimensional boundary space of $\partial\mathcal{B}(G)$ but a single point of $\partial_{\text{th}}\text{Teich}(R)$.

In contrast to the Bers boundary, the modular group extends so as to become a group of automorphisms (self-homeomorphisms) of $\mathfrak{T}(G) \cup \partial_{\text{th}}\mathfrak{T}(G)$. The orbit of each point of ∂_{th} is dense. In view of these properties, ∂_{th} is alternately referred to as the *geometric boundary*.

5.10 Exercises and explorations

5-1. *Hyperbolic Poisson integral formula* [Ahlfors 1981, §5.7]. A hyperbolically harmonic function $u : \mathbb{H}^3 \to \mathbb{R}$ is a function that vanishes under the Laplace–Beltrami operator $\Delta_h u = 0$. The Laplace–Beltrami operator in the ball model with $|\vec{x}| = r < 1$ is

$$\Delta_h u = \frac{(1-r^2)^2}{4}\Big(\Delta u + \frac{2r}{1-r^2}\frac{\partial u}{\partial r}\Big),$$

and in the upper half-space model $\{(z,t),\ t>0\}$ is

$$\Delta_h u = t^2\Big(\Delta u - \frac{1}{t}\frac{\partial u}{\partial t}\Big).$$

Here Δ denotes the euclidean laplacian. The Laplace–Beltrami operator has the property that for any isometry g,

$$\Delta_h(u \circ g) = (\Delta_h u)\circ g.$$

Thus u and $u \circ g$ are simultaneously hyperbolically harmonic.

We will use the ball model of $\mathbb{H}^3$ and denote spherical measure on $\mathbb{S}^2 \equiv \partial\mathbb{H}^3$ by $d\omega(\zeta)$. Suppose $f(\zeta)$ is a measurable function on $\mathbb{S}^2$ with $\iint_{\mathbb{S}^2} |f(\zeta)|\, d\omega(\zeta) < \infty$. The function

$$u(x) = u_f(x) = \frac{1}{4\pi}\iint_{\mathbb{S}^2}\left(\frac{1-|x|^2}{|\zeta - x|^2}\right)^2 f(\zeta)\, d\omega(\zeta), \quad x \in \mathbb{H}^3,$$

has the following properties [Ahlfors 1981, Chapter V].

(i) $u(x)$ is hyperbolically harmonic, and in particular real analytic, for $x \in \mathbb{H}^3$.
(ii) $u(x)$ has radial limits $f(\zeta)$ a.e.
(iii) If g is a Möbius transformation, then $u_f \circ g(x) = u_{f\circ g}(x)$.
(iv) In particular if $f \circ g(\zeta) = f(\zeta)$ for a Möbius transformation g and almost all $\zeta \in \mathbb{S}^2$ then $u \circ g(x) = u(x)$.

Note that the expression $(1-|z|^2)/|\zeta - z|^2$ is the Poisson kernel in the unit disk.

In the following theorem G is a kleinian group such that each component Ω_i of $\Omega(G)$ is a quasidisk with $\mathrm{Stab}(\Omega_i) = G_i$ and G is not quasifuchsian or a $\mathbb{Z}/2$ extension of one. Suppose the Möbius transformation T has the property that T sends the exterior of Ω_2 onto the interior of Ω_1.

Theorem 5.10.1 Existence of invariant embedded surfaces. *Let $\chi(\zeta)$ be the characteristic function of Ω_1, namely with value 1 for $\zeta \in \Omega_1$ and zero elsewhere. Denote by $u(x)$ the "harmonic measure" defined by the Poisson integral above with $f(\zeta) = \chi(\zeta)$. Choose $r > \frac{1}{2}$ such that the level surface $S = \{x \in \mathbb{H}^3 : u(x) = r\}$ is smooth*; *it is also embedded. Then $S \cap h(S) = \varnothing$ for all $h \in \langle G, TGT^{-1}\rangle$, $h \notin G_1$.*

Proof. See [Kapovich 2001, Lemma 4.102]. The surface S separates points of $\mathbb{H}^3$ with values $u(x) > r$ from points with $u(x) < r$. It is embedded because the gradient flow is orthogonal to S at all points. We know already that S is invariant under G_1.

Suppose for some $g \in G,\ g \notin G_1$ we had $y \in S \cap g(S)$. Recall that $u_\chi \circ g^{-1}(x) = u_{\chi\circ g^{-1}}(x)$ so $g(S)$ is the level surface $u_{\chi\circ g^{-1}}(x) = r$ for the characteristic function $\chi \circ g^{-1}$ of $g(\Omega_1)$. Adding the two Poisson integrals evaluated at y,

$$1 < 2r = \frac{1}{4\pi} \iint_{\Omega_1 \cup g(\Omega_1)} \left(\frac{1 - |y|^2}{|\zeta - y|^2} \right)^2 d\omega(\zeta). \tag{5.4}$$

But this is impossible since the regions Ω_1, $g(\Omega_1)$ are disjoint so the Poisson integral on the right represents the harmonic measure of their union. Its values must be strictly between 0 and 1 in $\mathbb{H}^2$.

The argument shows that g maps the level-r surface over Ω_1 onto the level-r surface over $g(\Omega_1)$. We have shown the totality of all such surfaces are mutually disjoint. Likewise the map T maps the level-r surface S_2 over Ω_2 to the level r-surface $T(S_2)$ over the complement Ω_1' of Ω_1. Now $T(S_2)$ is the level-$(1-r)$ surface over Ω_1. This can have no points in common with the level-r surface over Ω_1. Note that T maps the side of S_2 facing Ω_2 to the side of $T(S_2)$ facing away from Ω_1. We could take $r = \frac{1}{2}$ unless we wanted the surfaces to be smooth, then we can take r arbitrarily close to $\frac{1}{2}$.

There is one more case to check. Take a component $\Omega_3 \subset \Omega_1'$. Can its level-$r$ surface intersect $T(S_2)$? If y were an intersection point then the integral on the right side of Equation (5.4) would have the value $1 = r + (1-r)$. This could happen only if $\Omega_3 = \Omega_1'$ and G were quasifuchsian or an extension of a quasifuchsian by an element that interchanged the two components. We have assumed that this is not the case.

It follows that the orbit of S under the group $\langle G, TGT^{-1}\rangle$ is the union of mutually disjoint surfaces. $\square$

The bottom of the spectrum of eigenvalues. The bottom of the L^2-spectrum of the hyperbolic laplacian $-\Delta_h$ on the hyperbolic manifold $\mathcal{M} = \mathcal{M}(G)$ is given by

$$\lambda_0(\mathcal{M}) = \inf_{f \in C_c^\infty(\mathcal{M})} \frac{\int_{\mathcal{M}} |\nabla f|^2\, dV}{\int_{\mathcal{M}} |f|^2\, dV},$$

where the infimum is taken over all C^∞ functions with compact support, and dV is the volume element. Thus $\lambda_0(\mathcal{M}) = 0$ if $\mathcal{M}$ has finite volume (for then the constants are in the competition).

For any geometrically finite $\mathcal{M} = \mathcal{M}(G)$ with infinite volume and $\operatorname{Area} \partial\mathcal{C}(\mathcal{M}) = 2\pi|\chi(\partial\mathcal{C}(\mathcal{M}))|$,

$$\frac{K}{(\operatorname{Vol} \mathcal{C}_1(\mathcal{M}))^2} \le \lambda_0(\mathcal{M}) \le 4\pi \frac{\operatorname{Area} \partial\mathcal{C}(\mathcal{M})}{\operatorname{Vol} \mathcal{C}(\mathcal{M})}.$$

Here $K > 0$ is a universal constant and $\mathcal{C}_1(\mathcal{M})$ denotes the distance-1 neighborhood of the convex core $\mathcal{C}(\mathcal{M})$. The volume $\operatorname{Vol} \mathcal{C}_1(\mathcal{M})$ is finite if $\operatorname{Vol} \mathcal{C}(\mathcal{M}) < \infty$ and if for some $\delta > 0$, $\partial\mathcal{C}(\mathcal{M})$ contains no compressible curves of length $< \delta$—this condition prevents long thin waists with compressible cross section for which the 1-neighborhood will have large volume [Thurston 1979, Proposition 8.12.1]. Also $\mathrm{Vol}\,\mathcal{C}_1(\mathcal{M}) > 0$ if G is fuchsian. The right inequality is proved in [Canary 1992] and the left in [Burger and Canary 1994].

It is amazing that the lowest eigenvalue is precisely determined by the Hausdorff dimension $d(\mathcal{M})$ of the limit set of G. Making use of the information in Exercise 3-20, and bringing in the Tameness Theorem, we can state the situation as follows:

Theorem 5.10.2 [Sullivan 1987; Bishop and Jones 1997; Canary 1992]. *Suppose $\mathcal{M} = \mathcal{M}(G)$ has infinite volume and G is nonelementary.*

(*i*) $\lambda_0(\mathcal{M}) = 0$ *and* $d(\mathcal{M}) = 2$ *if and only if G is geometrically infinite.*
(*ii*) $\lambda_0(\mathcal{M}) = 1$ *if and only if G is geometrically finite with* $d(\mathcal{M}) \le 1$.
(*iii*) $0 < \lambda_0(\mathcal{M}) = d(\mathcal{M})(2 - d(\mathcal{M})) < 1$ *if and only if G is geometrically finite with* $1 < d(\mathcal{M}) < 2$.

In case (iii) there is an L^2 eigenfunction f_0 corresponding to $\lambda_0(\mathcal{M})$: $-\Delta_h f_0 = \lambda_0(\mathcal{M}) f_0$.

There is a global version of (iii). Suppose G is geometrically finite and nonelementary. Set $\lambda_0^*(G) = \sup \lambda_0(\mathcal{M})$, $d^*(G) = \inf d(\mathcal{M})$ as $\mathcal{M} = \mathcal{M}(H)$ ranges over $\mathfrak{R}_{\mathrm{disc}}(G)$. Then $\lambda_0^*(G) = d^*(G)(2 - d^*(G))$, provided $\mathcal{M}(G)$ is not a handlebody [Canary et al. 1999]. In addition the cases that $\lambda_0^*(G) = 1$ or $d^*(G) = 1$ are identified.

5-2. *The pinching estimate* [Bers 1970a; McMullen 1999]. The following estimate is frequently used in situations of pinching.

Theorem 5.10.3. *Suppose Ω is a simply connected component of the ordinary set $\Omega(H)$ corresponding to an incompressible component S of $\partial\mathcal{M}(H)$. Suppose $h \in$* $\mathrm{Stab}(\Omega)$ *is loxodromic and corresponds to the geodesics α in the hyperbolic metric on S and $\alpha_* \subset \mathcal{M}(H)$. Then*

$$\mathrm{Len}_{\mathcal{M}(H)}(\alpha_*) \;\le\; 2\,\mathrm{Len}_S(\alpha). \tag{5.5}$$

Suppose instead Ω_1, Ω_2 are the components of a quasifuchsian group H corresponding to the surfaces $S^t = \partial^{\mathrm{top}}\mathcal{M}(H)$, and $S_b = \partial_{\mathrm{bot}}\mathcal{M}(H)$. Denote by $\alpha_*, \alpha^t, \alpha_b$ the geodesics in the corresponding hyperbolic metrics. Then

$$\mathrm{Len}_{\mathcal{M}(H)}(\alpha_*) \;\le\; 2\min(\mathrm{Len}_{S_b}(\alpha_b), \mathrm{Len}_{S^t}(\alpha^t)). \tag{5.6}$$

Proof. The modulus of the annulus $A = \{z : r < |z| < R\}$ is defined as

$$M_A = \frac{\log(R/r)}{2\pi} = \frac{1}{\lambda(c_A)},$$

where $\lambda(c_A)$ is the extremal length of the free homotopy class of curves c_A separating the boundary components; see [Ahlfors 1973, Chapter 4]. The length of the shortest geodesic in c_A with respect to the hyperbolic metric of A is $\pi\lambda(c_A)$ (Exercise 2-2).

Suppose first we have a quasifuchsian group G with invariant components Ω_1, Ω_2. Assume that the loxodromic $g : z \mapsto kz$ is in G, $|k| > 1$. Then $A_i = \Omega_i/\langle g\rangle$ is conformally equivalent to an annulus, $i = 1, 2$. Let L_i denote the length of the geodesic in c_{A_i} in the hyperbolic metric in A_i. As above, $\lambda(c_{A_i}) = L_i/\pi$. These are the same as the lengths of the geodesics corresponding to g in each of the two surfaces $\partial\mathcal{M}(G)$.

On the other hand consider the torus $T = \mathbb{C}/\langle g\rangle$. The annular regions A_i are embedded in T and are disjoint there. The central curves of both A_1 and A_2 belong to the free homotopy class of curves c_T in T. A well known inequality, in particular in [Ahlfors 1973, Theorem 4.2], says that

$$\frac{1}{\lambda(c_T)} \geq M_{A_1} + M_{A_2} = \frac{1}{\lambda(c_{A_1})} + \frac{1}{\lambda(c_{A_2})} = \frac{\pi}{L_1} + \frac{\pi}{L_2}.$$

In particular, $\pi\lambda(c_T) \leq \min(L_1, L_2)$.

Calculate $\lambda(c_T)$ as follows. Set $\varphi = \arg k,\ 0 \leq \varphi < 2\pi$, and $\tau = \log k = \log|k| + \varphi i$. Consider the group $X = \langle z \mapsto z + 2\pi i,\ z \mapsto z + \tau\rangle$. A fundamental polygon is spanned by the vectors $(2\pi i, \tau)$. Its area is $2\pi|\tau|\cos\phi$, where $0 \leq \phi < \pi/2$ is a vertex angle; $\cos\phi = \pm\cos\varphi$.

The map $z = e^w$ sends $\mathbb{C}/X$ onto the torus T in such a way that $[0, 2\pi]$ is mapped to a circle and the line segments parallel to $[0, \tau]$ are mapped into the class c_T. The translation $w \mapsto w + \tau$ is sent to $z \mapsto kz$.

Now $\pm\cos\varphi = \operatorname{Re}\tau/|\tau|$. The rectangle of length $|\tau|$ and height $2\pi\operatorname{Re}\tau/|\tau|$ is foliated by the line segments parallel to $[0, \tau]$; it serves as a fundamental region for T. Conclude that $\lambda(c_T) = |\tau|^2/(2\pi\operatorname{Re}\tau)$.

For the quasifuchsian group we end up with

$$|\log k| \leq 2\min(L_1, L_2).$$

Now return to the hypothesis of Theorem 5.10.3. Here we use the annulus $A = \Omega/\langle h\rangle$ with $\mathrm{Len}_{\Omega/H}(h) = \pi\lambda(c_A)$. If h is conjugate to $z \mapsto kz$ (we can assume $|k| > 1$) then

$$\log|k| \leq |\log k| \leq \frac{2\pi}{M_A} = 2\pi\lambda(c_A) = 2\,\mathrm{Len}_{\Omega/\mathrm{Stab}(\Omega)}(h). \qquad \square$$

In particular, suppose $\mathcal{M}(H)$ has bounded geometry, that is, $\mathrm{Len}_{\mathcal{M}(H)}(\alpha_*) \geq \epsilon > 0$, for all geodesics $\alpha_* \subset \mathcal{M}(H)$. In the quasifuchsian case, the lengths of geodesics in both the top and bottom component are uniformly bounded below. If we can find a path out to the Bers boundary in a Bers slice, and if all manifolds along this path have uniformly bounded geometry with geodesic lengths $\geq \epsilon > 0$, then no pinching can occur. The projection of the path to the moduli space lies in a compact set.

Suppose instead that Ω is a nonsimply connected component of $\Omega(H)$ invariant under a function group G. Assume there is a positive lower bound for the length of all closed curves in Ω in the hyperbolic metric in Ω. According to [Canary 1991], there is a constant $\kappa > 0$ with the following property. If $c \subset S = \Omega/G$ is a closed geodesic in the hyperbolic metric on S, and c^* is the geodesic or a point representing

c in $\mathcal{M}(H)$, then in the respective hyperbolic metrics,

$$\mathrm{Len}_{\mathcal{M}(H)}(c^*) \le \kappa\, \mathrm{Len}_S(c).$$

5-3. *Pinching.* Given a geometrically finite, boundary incompressible manifold $\mathcal{M}(G)$ we will say that disjoint simple loops $\alpha, \beta \subset \partial\mathcal{M}(G)$ are *parallel* if neither one can be homotoped to a puncture (or to a point), yet the loops are freely homotopic in $\mathcal{M}(G)$. That is, they bound an annular region on $\partial\mathcal{M}(G)$, or they bound an essential cylinder in $\mathcal{M}(G)$. Another way of saying this is that if $A \in G$ is a primitive loxodromic associated with α, and $B \in G$ one associated with β then the cyclic groups $\langle A\rangle$ and $\langle B\rangle$ are conjugate in G. By a primitive loxodromic A we mean that, for some lift $\alpha^* \in \Omega(G)$, A is a generator of cyclic subgroup that maps α^* onto itself. We have often said more simply that $A \in G$ is associated with α.

The best general result about pinching is as follows: The setting is a geometrically finite group G without elliptics. Let $\alpha_1, \dots, \alpha_n$ be mutually disjoint and nonparallel simple loops on $\partial\mathcal{M}(G)$ that are represented by the *loxodromics* $A_1, \dots, A_N$ and their conjugacy classes in G.

Pinching Theorem [Ohshika 1998a]. *The manifold $\mathcal{M}(G)$ can be pinched along the loops $\{\alpha_i\}$ resulting in a geometrically finite manifold $\mathcal{M}(H)$.*

More precisely there is a sequence of points $\{\theta_n : G \to G_n\}$ in the deformation space $\mathfrak{T}(G)$ such that

(i) *$\{G_n\}$ converges algebraically and geometrically to the group $H = \lim \theta_n(G)$;*
(ii) *$\lim \theta_n(A_i) = A_i^*$ is parabolic for $1 \le i \le N$, and every new parabolic in H is in the conjugacy class of some $\langle A_i^*\rangle$;*
(iii) *the interior $\mathbb{H}^3/G$ is homeomorphic to $\mathbb{H}^3/H$; and*
(iv) *if $\alpha_{i_1}, \dots, \alpha_{i_k}$ lie in the component R_i of $\partial\mathcal{M}(G)$, the surface $R_i \setminus \bigcup_{1\le j\le k} \alpha_{i_j}$ is homeomorphic to a union of components of $\partial\mathcal{M}(H)$ such that each α_{i_j} determines a pair of punctures.*

The statement has been slightly modified from Ohshika's. His proof depends on Thurston's stronger version of his Compactness Theorem (page 206) and brings in techniques used in [Maskit 1983] to prove a weaker result. In any case the existence of the limit group H follows from the Hyperbolization Theorem.

Prove this for quasifuchsian groups by replacing the α_i by ever thicker annuli as in Exercise 4-8. The estimate (5.5) is needed to show that the sequences $\{\theta_n(A_i)\}$ converge to parabolics.

Hint: Remove from the original manifold $\mathcal{M}(G)$ the geodesics which are parallel to the initial loops α_i. Apply the Hyperbolization Theorem to get a manifold where these geodesics correspond to rank two cusps. Do Dehn surgery on solid cusp tori.

For a different approach, apply Theorem 3.11.3.

5-4. *Anosov mappings of a torus.* Show that those automorphisms of a torus onto itself that preserve the free homotopy class of a simple loop are exactly those elements $\{A\}$ of the modular group (see Exercises 2-5 and 2-9) with $\mathrm{tr}^2(A) = 4$. So if $\mathrm{tr}^2(A) > 4$ no simple loops are preserved. In the latter case, the automorphisms are called *Anosov mappings*. Prototypical examples of such A are $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}2&1\\1&1\end{smallmatrix}\right)$.

Hint: Work with the square lattice and the quotient torus

$$\mathbb{T}^2 = \mathbb{C}/\langle z \mapsto z+1,\ z \mapsto z+i\rangle.$$

Referring back to Exercise 2-5, show that an affine map that sends the square lattice onto itself has the form $A : (x, y) \mapsto (u, v)$ where

$$u = ax + by,\ y = cx + dy, \quad a, b, c, d \in \mathbb{Z},\ ad - bc = 1.$$

Simple loops on the square torus are those coming from lines through the origin of rational slope. Show that A preserves a simple loop on the quotient if and only if $\mathrm{tr}^2(A) = 4$, that is if there is only one eigenvalue $\lambda = \pm 1$, or if and only if the fixed point of A, if $\neq \infty$, is a rational number. Otherwise the projection to $\mathbb{T}^2$ of each of the lines of given irrational slope is dense and uniformly distributed in the torus by a famous classical theorem of Weyl. The projection of the family of parallel lines is called a *foliation* of $\mathbb{T}^2$. Find the eigenvalues and the eigenvectors. Identify the set of orthogonal lines that are preserved by A. On one set of lines, A is expanding by a factor exceeding one; these are called the *stable leaves*. On the orthogonal lines, A is contracting; these are called the *unstable leaves*. An Anosov map is characterized by having such a pair of transverse, invariant foliations.

On tori, the Anosov maps have no singularities, but on hyperbolic surfaces, the topology forces singularities.

5-5. *Dehn twists.* A Dehn twist in the round annulus $A_0 = \{z : 1 < |z| < u\}$ is the following mapping. Positively orient the boundary components a, b of A_0. Hold one fixed and twist A_0 by rotating the other by 2π in its positive direction. For example we can take (Figure 5.17)

$$\tau : re^{i\theta} \mapsto re^{i\theta s(r)}, \quad \text{where } 1 \leq r \leq u \text{ and } s(r) = \frac{(u - 2\pi) + (2\pi - 1)r}{u - 1}.$$

Now if $A \subset R$ is instead a neighborhood of a simple loop $\alpha \not\sim 1$ on some surface, we can bring over the twist from the round annulus to A. Extend the Dehn twist in A to all R by setting it equal to the identity on $R \setminus A$. This is a Dehn twist of R about

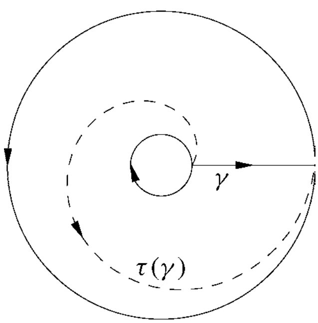


Fig. 5.17. A positive Dehn twist in an annulus.

α — we are only interested in its homotopy class. More precisely this is a Dehn twist of degree $+1$. Similarly we can define a Dehn twist of any positive or negative degree about α. A Dehn twist has no effect on the free homotopy class of a simple loop that does not make essential crossings with A.

For future use we also point out that a Dehn twist about a compressing curve α on $\partial\mathcal{M}(G)$ can be extended to a *twist map* of $\mathcal{M}(G)$, which is isotopic to the identity. First carry this out in the round model $\mathbb{D} \times [0, 1]$, via $(re^{2\pi i\theta}, t) \mapsto (re^{2\pi i(\theta+t)}, t)$. Then thicken to D^* an essential disk D bounded by α, so that $D^* \cap \partial\mathcal{M}(G) = A$. Bring over the twist to D^* and set it equal to the identity outside.

In the same vein we can speak of a Dehn twist about an essential cylinder C in $\mathcal{M}(G)$. Thicken C a bit to get the homeomorphic image C^* of the round $A_0^* = A_0 \times [-\delta, \delta]$. The Dehn twist in A_0 can be extended to A_0^* then C^*. Set it equal the identity outside C^*.

Suppose γ is a simple loop which has essential crossings with α, that is the number of crossings cannot be reduced within the free homotopy class of γ — this minimal number of crossings is the geometric intersection number $\iota(\gamma, \alpha)$. In the annular region A about α, we may assume $\gamma \cap A$ consists of a finite number of mutually disjoint segments running between the two boundary components. After A is twisted, each of the segments winds once around A but keeps the same endpoints. The collection remains mutually disjoint. The twisted segments can then be reattached to the part of γ outside A to become a new simple loop.

For example, consider the following construction on a genus 2 handlebody $\mathcal{M}(G)$ with boundary surface S. Let $\gamma \subset S$ be a simple dividing loop which in addition bounds a disk in the handlebody. Let $\alpha \subset S$ be simple nondividing loop which crosses γ twice and is not compressible in $\mathcal{M}(G)$. Let γ' be the result of applying a Dehn twist about α to γ. Show that γ' still divides S into two tori each with one boundary component but γ' is no longer compressible. In one of the components choose two simple loops A and B crossing each other exactly once. Show that the simple loop $[A, B]$ is freely homotopic to the boundary γ', and neither A or B is compressible.

5-6. *Surface automorphisms.* Suppose R is a surface of finite topological type. Let τ be an automorphism of R, that is, an orientation preserving homeomorphism of R onto itself. We are actually interested not so much in τ itself but in its equivalence class, also denoted by τ, under the homotopy relation $\sim$. These homotopy classes of automorphisms, as we saw on page 82, form the mapping class group $\mathfrak{M}(R)$. (One could use the isotopy equivalence relation instead, with the same results.)

Assume $\tau \not\sim \text{id}$. Thurston's classification of the automorphisms parallels the classification of Möbius transformations:

τ **has finite order** if for some integer, $\tau^m \sim \text{id}$;

τ **is reducible** if there is a set of mutually disjoint, unoriented, nontrivial simple loops $\mathcal{S} \subset R$, no two of which are parallel and no one of which is parallel to a boundary component or puncture, such that τ preserves the set $\mathcal{S}$ up to free homotopy of its elements;

τ **is pseudo-Anosov** if it has neither property (i) nor (ii).

Finite order elements. An element of finite order is analogous to an elliptic Möbius transformation. As mentioned on page 82, there was a longstanding conjecture of Nielsen that given a finite group (up to homotopy) of automorphisms τ there is a hyperbolic structure that can be put on R so that the group becomes a group of isometries of R onto itself. This was first proved in [Kerckhoff 1983], with different proofs in [Gabai 1992; Casson and Jungreis 1994].

Reducible mappings. The term "reducible" is used to suggest that if one cuts R along the curves of $\mathcal{S}$, some power τ^m ($m \geq 1$) maps each of the complementary regions onto itself and therefore can be analyzed in terms of its action on simpler surfaces. A Dehn twist is clearly reducible (with $m = 1$). Dehn twists are analogous to parabolic Möbius transformations. The "fixed point" of a Dehn twist (as well as of its inverse) is the geodesic lamination consisting of the geodesic representative of the simple loop. A set of Dehn twists about mutually disjoint loops generates a free abelian subgroup of $\mathfrak{M}(R)$. (The mapping class group itself is generated by $2g+1$ Dehn twists — see [Birman 1974].)

Pseudo-Anosov mappings. This is the generic case. For example, on a once-punctured torus choose simple closed geodesics a, b crossing exactly once. The composition of the Dehn twist about a with the Dehn twist about b is pseudo-Anosov.

A pseudo-Anosov τ does not preserve the homotopy class of any simple closed curve on the surface. Rather it preserves a pair of geodesic laminations which, like all geodesic laminations, are limits of simple closed curves. The situation is as follows:

Given the Riemann surface R on which τ is acting, there is a uniquely determined holomorphic quadratic differential with the following property: The quadratic differential determines a singular euclidean structure on R in terms of which τ is an affine map on neighborhoods of nonsingular points. In analogy to the torus case when one has eigenvectors, in the geometry of quadratic differentials there is a horizontal (stable) and vertical (unstable) foliation. These are determined by the horizontal and vertical trajectories — the pullback of the horizontal and vertical lines in $\mathbb{R}^2$ by the locally conformal map determined by the quadratic differential — which are invariant under τ. The quadratic differential is associated with the extremal Teichmüller map in the homotopy class of τ. See Exercise 5-23.

From this one obtains projective measured laminations $(\Lambda_\pm, \mu_\pm)$, uniquely associated with "fixed points" of τ, in analogy with the fixed points of a loxodromic. They are called the *attracting* and *repelling* (or *stable* and *unstable*) laminations associated with the pseudo-Anosov τ. We can get Λ_+ as $\lim \tau^{+n}(\gamma)$ and Λ_- as $\lim \tau^{-n}(\gamma)$, for any simple closed geodesic γ. The measures $\mu_\pm$ are obtained as the projectivized limit of the counting measures. In fact the "eigenvalues" $k^{\pm 1}$, $k > 1$, are such that $\iota(\tau(\gamma), \Lambda_\pm) = k^{\mp 1}\iota(\gamma, \Lambda_\pm)$ for any simple closed geodesic γ. The geodesic laminations $\Lambda_\pm$ fill up R in the sense that each component of $R \setminus (\Lambda_+ \cup \Lambda_-)$ is either a relatively compact simply connected region, or an annular region about a puncture containing its universal horodisk.

$\Lambda_\pm$ are *uniquely ergodic* in that, up to projective equivalence, there is only one transverse measure on the underlying geodesic lamination. In fact the corresponding projective measured laminations are dense in all projective measured laminations. In particular, $\Lambda_\pm$ are *arational* in that each complementary component is either a polygon or a polygon containing a single puncture. An arational lamination has the property that it is crossed by all closed geodesics and every other measured lamination with distinct support (Section 3.9).

Just like loxodromic Möbius transformations, two pseudo-Anosov automorphisms τ_1, τ_2 either have the same pair of fixed points or no common fixed points. In the latter case the subgroup $\langle \tau_1^m, \tau_2^n \rangle \subset \mathfrak{M}(R)$ is a free group, for sufficiently large m, n; see [Mosher $\geq$ 2007].

For the details of this theory in the topological case see [Thurston 1988; Fathi et al. 1979; Mosher $\geq$ 2007]. For a discussion of the analytic interpretation presented above, see [Marden and Strebel 1984; 1986].

5-7. *Twists and traces.* Suppose G is a fuchsian group without elliptics. Let T be a loxodromic element, and $A, B \in G$ any two other elements with fixed points distinct from those of T. Prove that $\text{trace}^2(AT^nB)$ goes to ∞ when n goes to $\pm\infty$. Conclude that the length of the geodesics on the quotient surface determined by the members of the corresponding sequence become infinite. (*Hint:* Work in UHP, take T to have fixed points $0, \infty$).

Prove that for $\{AT^nBT^{-n}\}$, the trace becomes infinite as well.

Suppose on the surface $R = \mathbb{H}^2/G$, γ is a simple nontrivial loop. Let τ denote the Dehn twist about γ. First assume that γ separates R. Let δ be a simple loop with geometric intersection number $\iota(\delta, \gamma) = 2$. The loop δ is homotopic to a composition $\alpha\beta$ where α, β lie in the two complementary components except for a common origin on γ. The twist τ^n sends $\alpha\beta$ to, say, $\alpha\gamma^n\beta\gamma^{-n}$. If instead γ does not separate, let δ be a simple loop with $\iota(\delta, \gamma) = 1$ and crossing γ once, at its origin. The result of applying τ^n is then to send δ to $\delta\gamma^n$. In either case, applied to the corresponding elements of G, show that the length of the corresponding geodesics becomes infinite as $n \to \pm\infty$.

5-8. *Nondensity of maximal cusps.* This is an example of Curt McMullen. Consider the quasifuchsian space of a geometrically finite group without parabolics. Choose a boundary group H without parabolics where the bottom end of $\mathcal{M}(H)$ is geometrically infinite and the top end is geometrically finite (a closed surface). This group cannot be approximated (algebraically) by maximal cusps. For suppose otherwise so that $H = \lim H_n$. Each maximal cusp H_n has the property that all the components of $\partial\mathcal{M}(H_n)$ are triply punctured spheres. Therefore the convex core of $\mathcal{M}(H_n)$ is bounded by a finite union of totally geodesic 3-punctured spheres, the number of components being independent of n. In view of Theorem 4.6.2 we know that all sequences that converge algebraically to H also converge geometrically. In particular this is true of $\{H_n\}$. On the other hand geometric convergence implies convergence of the convex hulls. This is impossible at the bottom end.

5-9. *More on collapsing mappings* (Thurston; see [Minsky 1994b]). Let $f : \mathbb{S}^2 \to \mathbb{S}^2$ be any continuous map whose restriction to the lower half-plane LHP is a homeomorphism. Set $\sigma = f(\mathbb{R} \cup \{\infty\})$. $y \in \sigma$ corresponds the closed set $C_y = \{f^{-1}(y)\} \subset \mathbb{S}^1 \equiv \mathbb{R} \cup \{\infty\}$. Let $\mathcal{C}_y$ denote the hyperbolic convex hull in LHP of C_y.

Verify that $\{\mathcal{C}_y,\ y \in \sigma\}$ is a partition of LHP into disjoint closed sets. (*Hint:* if $\mathcal{C}_y \cap \mathcal{C}_z \neq \varnothing$ there would be two pairs of points in $\mathcal{C}_y$ and $\mathcal{C}_z$ that separate each other on $\mathbb{S}^1$.)

Each $\mathcal{C}_y$ is a polygon with possibly infinitely many edges. The totality of edges forms a geodesic lamination Λ in LHP.

Now suppose in addition there is a fuchsian group G such that $f \circ g = \theta(g) \circ f$ for all $g \in G$ and θ is an isomorphism to a kleinian group $\theta(G)$. Then Λ is invariant under G and projects to a geodesic lamination on LHP$/G$.

5-10. *Nielsen transformations.* Suppose G is a free group on two generators X, Y. Nielsen proved that every automorphism of G is the composition of a finite number of automorphisms of the following elementary types.

(i) Interchange the two generators: $(X, Y) \to (Y, X)$.
(ii) Replace one generator by its inverse: $(X, Y) \to (X, Y^{-1})$.
(iii) Replace one generator by its product with the other: $(X, Y) \to (X, XY)$.

Thus if we start with a particular generating pair, we can systematically find every other generating pair.

This is applied to once punctured torus groups, where there is another relation that has to be maintained. By a generator pair of a once punctured torus group G we mean elements $X, Y \in G$ such that $G = \langle X, Y\rangle$ *and* $[X, Y]$ generates a parabolic subgroup.

The commutator requirement is preserved by the Nielsen transformations: Application of the first Nielsen transformation sends the commutator $[X, Y] = XYX^{-1}Y^{-1}$ to its inverse $[Y, X]$, the second sends it to $[X, Y^{-1}] = Y^{-1}[Y, X]Y$, and the third sends it to $X[X, Y]X^{-1}$. In any case the trace of the commutator remains -2.

Show using a cancellation argument that if both (X, Y) and (X, Z) are generator pairs in this sense then $Z = X^n Y$, modulo conjugation by some X^k.

For a quasifuchsian once punctured torus group G, set as usual $S^{\mathrm{top}} = \partial^{\mathrm{top}}\mathcal{M}(G)$, and $S_{\mathrm{bot}} = \partial_{\mathrm{bot}}\mathcal{M}(G)$. Suppose $\langle X^{\mathrm{top}}, Y^{\mathrm{top}}\rangle$ is a generator pair for $\pi_1(S^{\mathrm{top}})$ and the elements $X_{\mathrm{bot}}, Y_{\mathrm{bot}}$ are a generator pair for $\pi_1(S_{\mathrm{bot}})$. Assume these ordered generating pairs are not conjugate within G. Then we can pinch S^{top} and S_{bot} by requiring X^{top} and X_{bot} to become parabolic independently of each other. This results in a manifold whose boundary is two triply punctured spheres. Still, the two generator pairs are Nielsen transforms of each other.

Show that the Nielsen transformations generate the mapping class group for a once punctured torus.

5-11. *Geometric limits on the Bers boundary.* The point of this exercise is to identify the limit of a sequence $\{\tau^n(O)\}$ in a Bers slice $\mathcal{B}(R)$ as $n \to +\infty$. Here $O = (R, S_{\mathrm{bot}}, S^{\mathrm{top}})$ is a basepoint representing the manifold $\mathcal{M}(G)$ with G fuchsian; τ is a Dehn twist about a simple loop $\gamma \subset S_{\text{bot}}$. We will assume that γ divides R into two subsurfaces $S_{\text{bot}1}$, $S_{\text{bot}2}$. Let $\gamma_t \subset S^{\text{top}}$ be parallel to γ. It divides S^{top} into subsurfaces $S^{\text{top}1}$, $S^{\text{top}2}$ parallel to $S_{\text{bot}1}$, $S_{\text{bot}2}$.

The set of discontinuity $\Omega(G)$ has two components $\Omega^{\text{top}} = \text{UHP}$, $\Omega_{\text{bot}} = \text{LHP}$. The totality of lifts γ_t^* of γ_t divides Ω^{top} into countably many regions. Consider two adjacent regions $\Omega_{\text{top}1}$, $\Omega_{\text{top}2}$ where the first listed covers $S^{\text{top}1}$ and the second $S^{\text{top}2}$. Each is preserved by a subgroup G_1, G_2 of G, and G itself is the free product of these two subgroups, amalgamated over the common cyclic subgroup generated by the element preserving their common boundary. If $\Omega'_{\text{top}2}$ is another component adjacent to $\Omega_{\text{top}1}$ then its stabilizing group is a conjugate of G_2.

Denote the sequence of points in the Bers slice corresponding to $\{\tau_n(O)\}$ by

$$\{(S_{n,\text{bot}}, S^{n,\text{top}}; J_n = J \circ \tau^n)\}.$$

Neither the conformal type of the bottom or the top surface changes, but the topological relationship determined by J_n is causing increasing distortion which in the end will cause a fracture in the manifold. Each triple corresponds to a quasifuchsian group H_n which we are free to normalize so that for example 0, 1 are the fixed points of $\theta_n(g_1)$, and ∞ is the repelling fixed point of $\theta_n(g_2)$. Here $\theta_n : G \to H_n$ is the isomorphism determined by the conformal map F_n of LHP normalized to fix 0, 1, ∞. Its projection is a conformal map $f_n : S_{\text{bot}} \to S_{n,bot}$.

We may assume the sequence $\{F_n\}$ converges to a conformal mapping F of LHP and so $\{\theta_n(G) = H_n\}$ converges algebraically to $\theta : G \to H$. We may also assume that there is geometric convergence to some $H^* \supset H$.

What is happening on the top? For example take a simple loop $\delta \subset S_{\text{bot}}$ that is transverse to γ, crossing it exactly twice. We can write $\delta \sim \alpha\beta$, where α, β are disjoint except for a common basepoint on γ and α lies in $S_{\text{bot}1}$, β in $S_{\text{bot}2}$. In $S_{n,bot}$ we have $\delta_n = f_n(\delta) \sim \alpha_n\beta_n$ transverse to $\gamma_n = f_n(\gamma)$. With respect to the pairing J_n, δ_n is parallel in $\mathcal{M}(H_n)$ to $J_n(\delta_n) = \alpha_n {\gamma_n}^n \beta_n {\gamma_n}^{-n}$ (up to homotopy, also the use of $+n$ or $-n$ here depends on orientations). From Exercise 5-7 we know that the length of the corresponding geodesic on $S^{n,\text{top}}$ becomes infinite.

We conclude that the top surface is becoming pinched. That is if $g \in G$ is a loxodromic that corresponds to γ and γ_t, $g^* = \lim \theta_n(g)$ is parabolic. The boundary $\partial^{\text{top}}\mathcal{M}(H)$ is the union of two surfaces, homeomorphic to $S^{\text{top}1}$, $S^{\text{top}2}$ joined by a pair of punctures.

There exists a conformal map $\hat{F}_n : \text{UHP} \to \Omega_{n,\text{top}}$, but it *does not* induce θ_n there. Instead, we may assume that on $\Omega_{\text{top}1}$ $\hat{F}_n$ induces the restriction of θ_n to its stabilizer G_1. But on $\Omega_{\text{top}2}$, $\hat{F}_n$ induces the isomorphism $\theta_n(g^n)\theta_n(G_2)\theta_n(g^{-n})$ where $g \in G$ is a loxodromic that corresponds to γ_t. We may normalize $\hat{F}_n$ on $\Omega_{\text{top}1}$ and assume $\{\hat{F}_n\}$ converges to a conformal map $\hat{F}$ of UHP. The convergence forces $\{\theta_n(g^n)\}$ to converge to a Möbius transformation $h^* \neq \text{id}$, $h^* \in H^*$. That in turn forces the relation $h^* g^* h^{*-1} = g^*$ to hold. This relation in turn forces both g^*, h^* to be parabolic since they cannot both lie in a cyclic subgroup of H^*. Also h^* has the same fixed point as

g^*. So $\langle g^*, h^*\rangle$ is a rank two parabolic group and determines a solid cusp torus in $\mathcal{M}(H^*)$.

What is going on here? $\hat{F}$ maps UHP onto a component Ω^* of $\Omega(H^*)$ and induces an isomorphism of G onto the stabilizing subgroup G^* of Ω^*. The projection $\hat{F}_*$ maps S^{top} conformally onto $\hat{F}(\text{UHP})/G^*$. The image of $\Omega_{\text{top}1}$ in $\Omega(H)$ is bounded by “horocycles” associated with the parabolic fixed points coming from a component over $\partial^{\text{top}}\mathcal{M}(H)$. The rest of UHP is mapped into the corresponding “horodisks”. The extra parabolic h^*, maps the exterior of one “horodisk” onto the interior of its partner. Thus $\mathcal{M}(H^*)$ can be constructed from $\mathcal{M}(H)$ by the method of Exercise 4-18.

The geometric limit manifold $\mathcal{M}(H^*)$ is geometrically finite with the following structure. $\partial\mathcal{M}(H^*)$ has two boundary components, one conformally equivalent to $R = S_{\text{bot}}$, the other to S^{top}. $\mathcal{M}(H^*)$ is homomorphic to $(R \times [0, 1]) \setminus \{c\}$. Here $c \subset R \times \{1/2\}$ is a simple loop parallel to γ and γ_t. In $\mathcal{M}(H^*)$ it represents the solid cusp torus. We will deal with manifolds of this type in the next exercise.

Finally, the algebraic and geometric limits are independent of the subsequences used to attain them. This is a consequence of the Rigidity (or Isomorphism) Theorem 3.13.3.

The reader is invited to confirm the various assertions made. For another exposition with application to Riemann surface theory, see [Marden 1980]. Published details appear in the independent development in [Kerckhoff and Thurston 1990], where the result is applied to show that the extension of the mapping class group to the Bers boundary is not necessarily continuous.

Once we have the new groups H, H^* we can repeat the process with a new simple loop, and keep going until the loops selected determine a pants decomposition on $\partial\mathcal{M}(G)$.

5-12. *Piling up double cusps.* This construction is simplest for boundary cusps of a Bers slice in the once-punctured torus quasifuchsian space. For in this quasifuchsian space, a boundary cusp group H is such that $\partial\mathcal{M}(H)$ is the union either of a triply punctured sphere and a once-punctured torus (which we will here call a single cusp group), or of two triply punctured spheres (a double cusp group).

We can make an arbitrarily high pile of double cusp groups, where the top and/or bottom of the pile, if the pile is finite in that direction, is either a double cusp group or a single cusp group. It should suffice to illustrate the method in the simplest case.

Suppose H_1, H_2 are single cusp groups. We may arrange things so that the triply punctured sphere is the top component $S^{\text{top}1}$ of $\partial\mathcal{M}(H_1)$ and the bottom $S_{\text{bot}2}$ of $\partial\mathcal{M}(H_2)$. This means if we start with a fuchsian once-punctured torus group G, then there are orientation preserving homeomorphisms Φ_1, Φ_2 of the interiors $\mathbb{H}^3/G \to \mathbb{H}^3/H_i$ that take the top and bottom ends of $\mathbb{H}^3/G$ to the respectively labeled ends of $\mathbb{H}^3/H_i$.

Now $S^{\text{top}1}$ arises by pinching the top punctured torus along a simple closed curve and likewise pinching $S_{\text{bot}2}$. The pinching curves are represented on the other boundary components by simple loops $\gamma_1 \subset S_{\text{bot}1}$ and $\gamma_2 \subset S^{\text{top}2}$ respectively, parallel to the pinching curves. For our construction we must require that γ_1 and γ_2 are in the Φ_1, Φ_2 images of the *same* free homotopy class of $\mathbb{H}^3/G$.

Now of the three punctures on S^{top1} and S_{bot2}, one is spoken for as the puncture coming from the paired punctures on $\partial\mathcal{M}(G)$. We will refer to these as the basic punctures. The other two punctures will be called new punctures. If we draw a small circle about each of them, there is a solid pairing tube T_1, T_2 in each of $\mathcal{M}(H_1)$, $\mathcal{M}(H_2)$ that pairs them — its boundary is bounded by the two small circles. When we stack $\mathcal{M}(H_2)$ on top of $\mathcal{M}(H_1)$ so that the basic punctures are matched, the two solid pairing tubes will join up to form a solid cusp torus… and to determine a rank two cusp of the new manifold.

How do we do the construction so that the new manifold is hyperbolic? Represent S^{top1}, S_{bot2} as totally geodesic surfaces $S^{\text{top1}*}$, $S_{\text{bot2}}{}^*$ within the corresponding manifolds — just replace a disk in $\partial\mathbb{H}^3$ by the hyperbolic plane supported by its boundary and project. Let $P_1 \subset \mathbb{H}^3$ be a lift of $S^{\text{top1}*}$ and $\vec{n}_1$ the lift of an inner pointing normal. We may conjugate H_2 so that a lift P_2 of $S_{\text{bot2}}{}^*$ coincides with P_1 but $\vec{n}_2$ points to the opposite side of $P_1 = P_2$ as $\vec{n}_1$. We must also arrange things so that the respective subgroups preserving $P_1 = P_2$ are identical and the parabolic conjugacy class associated with the basic punctures coincide. Now $\langle H_1, H_2\rangle$ is discrete. How does the new rank two cusp arise?

Up in $\mathbb{H}^3$, given a lift P_1 of $S^{\text{top1}*}$, and a new parabolic α acting on P_1, there is another lift P_1' of $S^{\text{top1}*}$ which uniquely determined by the property that α preserves both — that the fixed point of α is the point of tangency of the boundaries of P_1 and P_1'. Choose an inner normal vector $\vec{n}_1'$ to P_1'. Correspondingly for $S_{\text{bot2}}{}^*$ there is another lift P_2' that is also preserved by α; here $P_2' \neq P_1'$. Choose an inner normal vector $\vec{n}_2'$. When H_1 and H_2 are joined across $P_1 = P_2$, consider the associated planes P_1' and P_2' which also share the fixed point of α on their boundaries. To complete the combination of the two groups, we must add another parabolic β that maps the side of P_1' containing $\vec{n}_1'$ onto the side of P_2' opposite that determined by $\vec{n}_2'$, and satisfying

$$\beta\alpha\beta^{-1} = \alpha, \quad \beta(P_1') = P_2', \quad \beta\text{Stab}(P_1')\beta^{-1} = \text{Stab}(P_2').$$

These conditions uniquely determine β. In essence we are making a construction as in Exercise 4-19. We end up with the group $H = \langle H_1, H_2, \beta\rangle$.

Verify that the hyperbolic construction works, and that the resulting manifold $\mathcal{M}(H)$ has the following properties:

(i) $\partial\mathcal{M}(H)$ has two components, one conformally equivalent to the bottom component of $\partial\mathcal{M}(H_1)$, the other to the top component of $\partial\mathcal{M}(H_2)$.

(ii) In the interior of $\mathcal{M}(H)$ there is a "singular" totally geodesic surface representing the thrice punctured sphere.

(iii) $\mathcal{M}(H)$ is homeomorphic to $(S \times [0, 1]) \setminus c$ where c is a circle homotopic to the representative of the pinching curve on each of the two boundary components, and S is a once-punctured torus.

Actually we made no essential use of the fact that the bottom component of $\partial\mathcal{M}(H_1)$ and top of $\partial\mathcal{M}(H_2)$ remained a once punctured torus. H_1 and H_2 could as well have been double cusp groups. In this case we can continue the process and build an arbitrarily large pile of manifolds. This is the process required to directly construct the groups representing geometric limits in quasifuchsian space, as described in Section 5.9.

5-13. *Shuffling a rolodex.* This and the next exercise is a report on [Anderson and Canary 1996a]. First we will construct the basic manifold.

Start with a solid torus T and its core curve c. Fix a finite system of mutually disjoint, parallel simple loops $\{\gamma_k\}$ on ∂T which are not contractible in T. Correspondingly fix a collection of surfaces $\{S_k\}$, each of some genus $g_k \geq 1$ and with a single boundary component. For greater effect assume the genera g_i are all different. Slightly thicken each S_k to obtain the compact manifolds $\{S_k \times [-\epsilon, +\epsilon]\}$. The boundary of each contains the annulus $\partial S_k \times [-\epsilon, \epsilon]$.

Attach $S_k \times [-\epsilon, \epsilon]$ by gluing $\partial S_k \times [-\epsilon, \epsilon]$ to a thin neighborhood of γ_k. The resulting manifold M is orientable and compact. By "rearranging the pages"—taking a noncyclic permutation τ of $\{S_k\}$, we get another manifold M_τ which is homotopy equivalent but not homeomorphic to M. The manifolds M_τ have a hyperbolic structure. For more details see Exercise 5-13.

By the Hyperbolization Theorem (page 324) we can write $M = \mathcal{M}(G)$, for a kleinian group G and likewise $M_\tau = \mathcal{M}(G_\tau)$. The original solid torus becomes a tubular neighborhood about the core geodesic c. So this construction results in a multitude (depending on the number of pages chosen) of hyperbolic manifolds homotopy equivalent but not homeomorphic to $\mathcal{M}(G)$ or to each other.

Let $\widehat{M}$ denote the result of removing from M the core curve c of T. Note that $T \setminus \{c\} \subset \widehat{M}$ has the structure of a solid cusp torus. By applying the Hyperbolization Theorem we can assume that $\widehat{M} \cong \mathcal{M}(H)$ for a geometrically finite H with now a rank two cusp.

To get bumping, do Dehn surgery (Section 4.9) on the rank two cusp of $\mathcal{M}(H)$. Take a cusp torus parallel to ∂T, a meridian α which is contractible in T, and a longitude β which is parallel to c. Set $\delta_n = \alpha + n\beta$, $0 \leq n$. After Dehn surgery δ_n becomes a meridian on the solid torus that replaces the solid cusp torus. The resulting manifold M_n also has a hyperbolic structure $\mathcal{M}(H_n)$ and is homeomorphic to $\mathcal{M}(G)$. We claim that $\{H_n\}$ converges algebraically to a geometrically finite group G^* with a rank one parabolic and geometrically to $H \supset G^*$ in analogy with Section 4.9.

This is seen by considering the sequence of representations $\psi_n : d_n \circ \iota$ of G where d_n is the homeomorphism $H \to H_n$ given by Dehn surgery, and $\iota : \pi_1(M) \to \pi_1(\widehat{M})$ is the inclusion. The image group is H_n and ψ_n is induced by a homeomorphism between the manifolds $\mathcal{M}(G)$ and $\mathcal{M}(H_n)$.

In fact we will explicitly construct G^* and H in the next exercise. It will turn out that in H the $\{S_k\}$ are once punctured surfaces and the system $T \setminus \left(\bigcup \partial S_k \times (-\epsilon, \epsilon)\right)$ appears as pairing tubes, pairing successive punctures.

Next, choose a noncyclic permutation τ of $(1, 2, \ldots, k)$. Using the permutation of indices given by τ, build M_τ on T as M was built. As pointed out earlier, the manifolds M, M_τ have isomorphic fundamental groups so they are homotopy equivalent, but they are not homeomorphic. There is an isomorphism $\theta_\tau : G_\tau \to G$.

The hardest part is to construct an immersion $f_\tau : M_\tau \to \widehat{M}$ and equally $\mathcal{M}(G_\tau) \to \mathcal{M}(H)$ such that on the level of fundamental groups f_τ has the properties (i) it determines an isomorphism of G_τ onto a geometrically finite subgroup of H, i.e., $\pi_1(M_\tau) \to \pi_1(\widehat{M})$, and (ii) it determines an isomorphism $\theta_{n,\tau} : G_\tau \to H_n$, i.e., $\pi_1(M_\tau) \to \pi_1(\widehat{M}) \to \pi_1(M_n)$.

The sequence of isomorphisms $\{\theta_{n,\tau}\}$ converges algebraically to the isomorphism $\theta_\tau : G_\tau \to G^* \subset H$ and the sequence of groups geometrically to H itself. Thus the group G^* is a boundary group of the deformation space of G and of the deformation space of G_τ. These groups represent the nonhomeomorphic manifolds $\mathcal{M}(G)$ and $\mathcal{M}(G_\tau)$. The two deformation spaces bump at G^*.

Manifolds of the type constructed above exhibit another interesting property (lest one believes such a phenomenon does not occur!): An example of a hyperbolic manifold which has a simple geodesic γ which is not freely homotopic to curve in the boundary, yet γ^n is freely homotopic to a simple loop in the boundary.

Instead of doing $(1, n)$-Dehn surgery on $\mathcal{M}(H)$ do $(n, 1)$-Dehn surgery. That results in a manifold $\mathcal{M}(H_n')$ in which $\beta\alpha^n$ is homotopic to a point. That is, $\beta \sim \alpha^{-n}$. Now in $\mathcal{M}(H_n')$, α cannot be homotoped into $\partial\mathcal{M}(H_n')$ (or to a point). Yet β can be homotoped into the boundary — it is parallel to the central curves of the annuli we used.

5-14. *Constructing a rolodex.* In this exercise we will explicitly construct the rolodex used in Exercise 5-13. Again we closely follow [Anderson and Canary 1996a].

We used above $k \geq 3$ surfaces $\{S_k\}$ with one boundary component and distinct genera $\{g_i \geq 1\}$. We will now assume the surfaces are closed Riemann surfaces each with one puncture.

Uniformize each of the surfaces by a fuchsian group $\{\Gamma_k\}$ acting in the upper and lower half-plane so normalized so that ∞ is a parabolic fixed point and $z \mapsto z+1$ generates the rank one parabolic group at that point. By the universal disk property, the horizontal strip $\sigma = \{z \in \mathbb{C} : -1-\varepsilon < \operatorname{Im} z < 1+\varepsilon\}$, $\varepsilon > 0$, has the following property. For any element $g \neq \mathrm{id}$ of any group Γ_k, $g(\mathbb{C} \setminus \sigma) \subset \sigma$, unless g is in the rank one parabolic group at ∞. Moreover, the vertical slab σ^* in upper half-space over σ has the property that $\sigma^*/\Gamma_k \cong S_k \times [0, 1]$. Let σ' be the result of truncating σ^* at height $1+\varepsilon$. By the Universal Horoball Theorem, we have that $g(\mathbb{C} \cup \mathbb{H}^3 \setminus \sigma^*) \subset \sigma^*$ for any element g of any Γ_i, provided g does not fix ∞. Instead of lining the surfaces up on a solid torus as above, we will line them up in vertical translates of σ.

Next choose $\mu > 2k(1+\varepsilon)$. Conjugate each Γ_j by a vertical translation $z \mapsto z + a_j i$ so that the horizontal strips σ_j for the resulting groups all lie in $\{z : 0 < \operatorname{Im} z < \mu\}$, with mutually disjoint closures and with the order $\sigma_1, \sigma_2, \ldots, \sigma_k$ as $\operatorname{Im} z$ increases from $\operatorname{Im} z = 0$ to $\operatorname{Im} z = \mu$. Denote the conjugated groups by $\{\Gamma_k'\}$.

We claim that the group $G^* = \langle \Gamma'_1, \Gamma'_2, \ldots, \Gamma'_k \rangle$ is a kleinian group such that $\mathcal{M}(G^*)$ with the interior of a solid pairing tube removed is homeomorphic to the complement of the interior of T in the manifold M constructed above. In G^* there is only one parabolic conjugacy class, namely that generated by $z \mapsto z+1$, while if $g(\infty) \neq \infty$, $g \in \Gamma_j$ maps the exterior of σ'_j into σ'_j.

Let $U(z) = z + \mu i$ and set $H = \langle G^*, U \rangle$. We claim that $\mathcal{M}(H)$ is homeomorphic to $\widehat{M}$ constructed above. H is geometrically finite with a rank two parabolic group at ∞.

Finally we construct a hyperbolic structure for M_τ. To do that we have to rearrange the strips $\{S_i\}$ to have the new ordering dictated by τ. A simple way of doing that is as follows:

$$H_\tau = \langle U\Gamma_{\tau(1)}U^{-1}, U^2\Gamma_{\tau(2)}U^{-2}, \ldots, U^k\Gamma_{\tau(k)}U^{-k} \rangle.$$

5-15. *The curve complex.* In the hands of Howard Masur and Yair Minsky [1999; 2000] this has proved to be an essential tool in analyzing the short geodesics near the ends of hyperbolic manifolds. The definition of the complex was originally given by Bill Harvey [1981]. Let Σ denote a compact surface of genus $g \geq 0$ and $n \geq 0$ boundary components.

The simplicial complex $\mathcal{C}(\Sigma)$ is defined as follows:

(i) The vertices V are the simple closed curves on Σ that cannot be homotoped into a boundary component or to a point.
(ii) Two vertices $v_1, v_2 \in V$ are joined by an edge e if v_1 and v_2 correspond to disjoint simple loops, not in the same free homotopy class.
(iii) More generally, a finite set $\sigma \in V$ will bound a simplex in $\mathcal{C}(\Sigma)$ if the elements of σ can be realized by mutually disjoint, nonparallel, simple loops.

Special cases are:

- $g = 0,\ n \leq 3$. Then $V = \varnothing$; there are no simple closed curves.
- $g = 0, n = 4$, or $g = 1, n = 0, 1$. Then $\mathcal{C}(\Sigma) = V$; there are no edges.

In the cases $g = 0, n = 4$ or $g = 1, n = 1$ the edges must be defined slightly differently: Vertices v_1, v_2 are to be connected by an edge if the simple loops representing the vertices have intersection number 2 in the first case and 1 in the second case. Then $\mathcal{C}(\Sigma)$ is isomorphic to the Farey graph in $\mathbb{H}^2$ (Exercise 2-9). For the remainder of the discussion we will exclude these cases.

The complex $\mathcal{C}(\Sigma)$ is connected and of dimension $(3g+n-4)$ meaning the largest simplices have $(3g+n-3)$ vertices. For example, if $g = 2, n = 0$ there are at most three simple, mutually disjoint, nonparallel, nontrivial simple loops. Thus the largest simplices in $\mathcal{C}(\Sigma)$ are triangles.

The curve complex is connected. However it is not locally finite. For suppose α, β are simple loops which cross once. They cannot be the endpoints of an edge. Rather, there are infinitely many distinct two-edge paths connecting them.

The mapping class group of Σ, namely the group of homotopy classes of homeomorphisms of Σ onto itself, acts on $\mathcal{C}(\Sigma)$.

The 1-skeleton $\mathcal{G}(\Sigma)$ of $\mathcal{C}(\Sigma)$ is a connected graph. With the assignment of length one to each edge, $\mathcal{G}(\Sigma)$ becomes a metric space. In [Masur and Minsky 1999] it is proved that $\mathcal{G}(\Sigma)$ is a Gromov hyperbolic space! See Exercise 2-17. Moreover Klarreich (unpublished) proved that the Gromov boundary of the curve complex can be identified with the space of ending laminations associated with Σ.

5-16. *The Masur domain.* Compressible ends have been the most difficult to analyze. A principal tool in the analysis is the Masur domain (see [Canary 1993; Kleineidam and Souto 2003]) first defined for compression bodies. For the extension to all geometrically finite manifolds see [Lecuire 2004c]. For this exercise, all kleinian groups are assumed to have *no parabolics* (although rank two parabolics could be allowed if the interior of solid cusp tori were removed from the manifolds to make them compact).

First we make some comments about compressible boundary components.

Suppose $\mathcal{M}(G)$ is compact (no parabolics). Assume that f is a quasiconformal automorphism $\mathcal{M}(G) \to \mathcal{M}(G)$ that fixes a basepoint $O \in \operatorname{Int}\mathcal{M}(G)$. Then the restriction of f to $\operatorname{Int}\mathcal{M}(G)$ induces the identity automorphism of $\pi_1(\operatorname{Int}\mathcal{M}(G)$ if and only if f is isotopic in $\operatorname{Int}\mathcal{M}(G)$ to a composition of Dehn twists about compressing curves [McCullough and Miller 1986, Theorem 6.2.1] (recall from Exercise 5-5 that a twist can be extended to $\mathcal{M}(G)$ and is isotopic to the identity there). Presumably some form of this result can be applied if rank one cusps are present as well (rank two cusps are not involved).

Suppose now $F : \Omega(G) \to \Omega(H)$ is a quasiconformal map of a geometrically finite manifold $\mathcal{M}(G)$ that induces an isomorphism $\varphi : G \to H$. By Theorem 3.7.3, we may assume that F extends to a quasiconformal mapping of $\mathbb{S}^2$ and projects to a quasiconformal map $f : \mathcal{M}(G) \to \mathcal{M}(H)$ that also induces the isomorphism $\varphi : \pi_1(\mathcal{M}(G)) \to \pi_1(\mathcal{M}(H))$. If $\varphi = \mathrm{id}$, necessarily F pointwise fixes $\Lambda(G)$.

We will now restrict our attention to a compact compression body (function group) $\mathcal{M}(G)$ and to its one compressible boundary component $S \subset \partial\mathcal{M}(G)$.

In general, there are infinitely many free homotopy classes of compressible loops — the exceptions occur when there is essentially only one compressing disk in $\mathcal{M}(H)$. A quasiconformal $f : \partial\mathcal{M}(G) \to \mathcal{M}(G)$, extends to a quasiconformal map of $\mathcal{M}(G)$ if and only if f preserves the set of free homotopy classes of compression curves [McCullough and Miller 1986]. As mentioned above, an automorphism f of $\mathcal{M}(G)$ that fixes a basepoint O in its interior induces the identity automorphism of the group $\pi_1(\mathcal{M}(G); O)$ if and only if it is isotopic (with O fixed) to a composition of Dehn twists about compression loops. If f is orientation reversing, $\mathcal{M}(G)$ is a handlebody (Luft's Theorem: see [McCullough and Miller 1986, Theorem 5.3.1]).

The compression body is called *small* if there is only one compression loop up to free homotopy, that is if for a compression disk D, $\mathcal{M}(G) \setminus D$ is one or two manifolds of the form $S_0 \times [0, 1]$ where S_0 is an incompressible component of $\partial\mathcal{M}(G)$.

In Section 3.9 we introduced the measured lamination space $\mathcal{ML}(S)$ and projective measured lamination space $\mathcal{PML}(S)$, using the hyperbolic metric on S (although that is not necessary). A finite leaved geodesic lamination is called compressible if each of its leaves is compressible. Let $C(S)$ be the set of projective classes of compressible finite leaved measured laminations with atomic measures on the leaves. Let $\mathcal{C}(S)$ denote the closure of $C(S)$ in $\mathcal{PML}(S)$.

Suppose the compact compression body $\mathcal{M}(G)$ is not small. While $\ell \in \mathcal{C}(S)$ if it is just a compressing curve, the following elements of $\mathcal{C}(S)$ are perhaps surprising. The geodesic lamination Λ on the compressible boundary component S is the support of a lamination in $\mathcal{C}(S)$ provided one of the following hold (Otal's thesis; see [Kleineidam and Souto 2003]):

(i) Λ is the union of two disjoint, simple loops which are not parallel on S but which bound an essential cylinder C within $\mathcal{M}(G)$, or
(ii) Λ is a simple loop which homotopic to α^k, for a loop $\alpha \in \pi_1(\mathcal{M}(G))$ and $|k| \geq 2$, or
(iii) Λ is a simple loop freely homotopic to a simple loop on $\partial\mathcal{M}(G) \setminus S$.
(iv) [Kleineidam and Souto 2003, Lemma 3.6] Λ is a minimal lamination for which $S \setminus \Lambda$ is compressible.

The first statement follows from the fact that since $\mathcal{M}(G)$ is not “small”, there is a compressing loop $\gamma \subset S$ which is transverse to both components of $C \cap S$ [Kleineidam and Souto 2003]. Let $\tau : \mathcal{M}(G) \to \mathcal{M}(G)$ be the Dehn twist about C (see Exercise 5-6). The the sequence of compressing loops $\{\tau^k(\gamma)\}$ converge to a measured lamination with support in Λ.

If M is not small the *Masur domain* of the compressible boundary component S is defined to be

$$\mathcal{O}(S) = \{\mu \in \mathcal{PML} : \iota(\lambda, \mu) > 0 \text{ for all } \lambda \in \mathcal{C}(S)\}.$$

If instead M is small then $\mathcal{O}(S)$ is defined to consist of $\mu \in \mathcal{PML}$ for which $\iota(\lambda, \mu) > 0$ for all those $\lambda \in \mathcal{PML}(S)$ for which there exists $\nu \in \mathcal{C}(S)$ with $\iota(\mu, \nu) = 0$.

The Masur domain is open in $\mathcal{PML}$. If the support of $\mu \in \mathcal{O}(S)$ is a finite number of simple geodesics, then every component of $S \setminus \mu$ is *incompressible* and even *acylindrical*. When the support of μ is a simple geodesic γ, γ is transverse to every simple compressing geodesic. Note that two simple compressing geodesics may well cross each other.

Denote by $\mathrm{Mod}(S)$ the group of orientation preserving automorphisms of S that extend to diffeomorphisms of $\mathcal{M}(G)$ and $\mathrm{Mod}_0(S)$ the subgroup whose extensions are homotopic to the identity. The group $\mathrm{Mod}(S)$ acts on $\mathcal{PML}(S)$ and $\mathcal{C}(S)$ is its limit set. $\mathrm{Mod}(S)$ also acts properly discontinuously on the Masur domain. Suppose the supports of both μ_1, $\mu_2 \in \mathcal{O}(S)$ are collections of simple geodesics on S. Then if the components of the supports are respectively freely homotopic within the compression body M, there is a element $h \in \mathrm{Mod}(S)$ with $h(\mu_1) = \mu_2$ (Otal; see [Canary 1993]).

For further details see [Masur 1986; Kerckhoff 1990; Canary 1993; Kleineidam and Souto 2003].

Regarding compressible ends, we have the following restricted but striking result; see also [Kleineidam and Souto 2002; Ohshika 2005, Theorem 4.1].

Theorem 5.10.4 [Kleineidam and Souto 2003, Corollary 1.2]. *Suppose H has no parabolics and is not a free group. Then $\mathcal{M}(H)$ is tame if and only if for every geometrically infinite, compressible end E of $\mathcal{M}(H)$, there is a Masur domain lamination on ∂E that is not realized in $\mathcal{M}(H)$.*

Of course we now know every $\mathcal{M}(H)$ is tame. Recall that an end can be represented by a boundary component of the compact core. The case of incompressible ends is covered by Bonahon's criteria (page 253).

5-17. *Diskbusting curves; Canary's trick.* Here we are only concerned with manifolds $\mathcal{M}(G)$ that have nontrivial splittings into free products $\pi_1(\mathcal{M}(G)) = A * B$. A *diskbusting curve* $\sigma \in \pi_1(\mathcal{M}(G))$ is one that is not contained in either factor of any free product decomposition of $\pi_1(\mathcal{M}(G)) \cong G$. Thus if $\pi_1(\mathcal{M}(G)) = \langle g_1, g_2 \ldots g_N \rangle$ is a free group on N generators, then the element $g = g_1^2 g_2^2 \cdots g_N^2$ is diskbusting as it is the relator of a closed, nonorientable surface with an N-generator fundamental group.

If G is finitely generated but $\mathcal{M}(G)$ is not known to be tame, then $\pi_1(M(G) \setminus \sigma)$ may not be finitely generated. This issue is the root of much trouble.

In the case of compact, hyperbolizable manifolds M^3 the algebraic definition we have just given is equivalent to the geometric definition: A curve $\sigma \in \pi_1(M^3)$ is *diskbusting* if any curve $\sigma' \in M^3$ freely homotopic to σ intersects every compressing disk in M^3. For according to Proposition 3.7.1 every free product decomposition of M^3 is generated by a compressing disk. Thus a simple loop on the boundary of a compression body M^3 such that it, with its counting measure, lies in the Masur domain is diskbusting [Canary 1993, Proposition 3.4]. In fact, there is a countable collection of them, no two of which are freely homotopic in M^3.

If β is a curve which is not based at the basepoint of $\pi_1(\mathcal{M}(G))$, we will say β is diskbusting if it is freely homotopic to a diskbusting curve based at the basepoint. Thus a diskbusting curve is one for which there exists an auxiliary arc x from its basepoint to the basepoint of $\pi_1(\mathcal{M}(G))$ such that $x\beta x^{-1}$ is diskbusting.

The exists an infinite collection of diskbusting curves on the compressible boundary component S of a compression body $\mathcal{M}(G)$, no pair being freely homotopic in $\mathcal{M}(G)$. Moreover, if the support of $\sigma \in \mathcal{O}(S)$ (Exercise 5-16) is a simple geodesic on S, then σ is diskbusting [Canary 1993, Proposition 3.4 and Corollary 3.5].

Now a diskbusting curve β, or any curve freely homotopic to it, has the property that if D is any compressing disk based on the boundary, then β intersects D. Thus if $\mathcal{M}(G)$ is geometrically finite with compressible boundary, or if we are in a relative compact core C, then $\mathcal{M}(G) \setminus \beta$ or $C \setminus \beta$, is incompressible. We can take diskbusting curves to be geodesics, but we cannot be sure they are simple.

A union of mutually disjoint simple closed curves is called a *diskbusting link* if for every free product decomposition, at least one of the components of the link is not contained in either factor of the decomposition. We can add simple curves to the link and its diskbusting role will not change. Such a diskbusting link can be taken to be composed on geodesics; they will be mutually disjoint but be cannot be sure they are simple.

Now suppose σ is diskbusting link in the interior of $\mathcal{M}(G)$ which is homologous to zero. For example σ might lie on the boundary of a compact core as in Exercise 5-16. We can take σ to consist of closed geodesics. Suppose they are all simple. According to [Canary 1993, Lemma 3.1], σ bounds an embedded oriented surface $\Sigma \subset \mathcal{M}(G)$.

Next we will construct a new 3-manifold $\widehat{M}$ which is a two sheeted branched cover of $\mathcal{M}(G)$, branched over σ. This is easily effected by cutting open $\mathcal{M}(G)$ along Σ and designating the two sides of the cut by $\Sigma_{\pm}$. Set $M = \mathcal{M}(G) \setminus \Sigma$. Take two copies of M and identify the $\pm$ sides of Σ on one copy to the $\mp$ copies of Σ on the other.

Lemma 5.10.5 [Canary 1993, 5.2]. *$\partial \widehat{M}$ has incompressible boundary.*

This is a great advantage for analyzing ends. The bad news is that we have introduced cone axes. In Exercise 5-17, it is shown how to smooth these out, and also how to deal with the situation that one or more of the geodesics in the link is not simple.

Diskbusting links are used to deal with compressible ends. In studying the corresponding branched cover, the *engulfing property* of Brin and Thickstun is used (see [Agol 2004], [Myers 2005]): Suppose X is a compact, connected submanifold of an orientable, irreducible 3-manifold M^3 (for example, as in Exercise 5-18) with no compact complementary components. There exists an open (not necessarily properly embedded) submanifold Y containing X, uniquely determined in M^3 up to isotopy fixing X by the following properties.

(i) Y has no compact complementary components,
(ii) Y has a regular exhaustion $\{Y_n\}$ such that ∂Y_n is incompressible in the complement of X,
(iii) Given compact submanifold Z with $X \subset \operatorname{Int} Z \subset Z \subset M^3$ and ∂Z incompressible in the complement of X, then Z can be isotoped into Y with the isotopy fixing X.

The submanifold Y is called an *end reduction* at X.

Most directly applicable is the following:

Theorem 5.10.6 [Myers 2005]. *Let $\alpha \subset M^3$ be an algebraically diskbusting link with $X = N(\alpha)$ a thin tubular closed neighborhood of α. If V is an end reduction, then the inclusion $\iota : \pi_1(V) \hookrightarrow \pi_1(M^3)$ induces an isomorphism $\pi_1(V) \to \pi_1(M^3)$, and the inclusion $\pi_1(V \setminus \alpha) \hookrightarrow \pi_1(M^3 \setminus \alpha)$ is an injection.*

5-18. *Pinched negative curvature manifolds.* It has been quite useful to eliminate pesky cusps, inconvenient cone singularities, and self-intersections of geodesics by locally changing the constant negative curvature hyperbolic metric (singular on the cone axes) to a complete riemannian metric of *pinched negative curvature* (PNC): the sectional curvature lies between two constants $-b^2 < -a^2 < 0$. For PNC-manifolds, many of the qualitative properties of hyperbolic manifolds remain true — see [Canary 1993] and additional references listed there. A general reference is [Ballmann et al. 1985].

The following result is attributed to Gromov and Thurston and proved in [Bleiler and Hodgson 1996].

Theorem 5.10.7 (The "2π Lemma" or "Theorem A"). *Let V be a solid torus with a hyperbolic metric near ∂V so that ∂V is the quotient of a horosphere. The hyperbolic metric can be extended to a PNC metric in V provided that the length of the euclidean geodesic on ∂V serving as meridian has length $> 2\pi$.*

The necessity of 2π follows from the Gauss–Bonnet formula Equation (1.3) applied to a geodesic meridian, which inherits curvature $+1$ from the horosphere Exercise 3-33, and bounds a disk of negative curvature in V.

Agol [2004] recognized a very useful generalization to be used in the frequent cases when the 2π condition is not satisfied. I thank Juan Souto for pointing this out.

Theorem 5.10.8 (The "$2\pi/k$ Lemma" or "Theorem A^+"). *Let V be a solid torus or an infinite cylinder with a hyperbolic metric near ∂V so that ∂V is the quotient of a horosphere. Find $k \geq 1$ such that the length of the euclidean geodesic on ∂V serving as meridian has length $\geq 2\pi/k$. Then the hyperbolic metric near ∂V can be extended so that the solid torus or cylinder V becomes a PNC $2\pi/k$ orbifold* (*when $k \geq 2$*).

Consequently rank one or rank two cusps can always be eliminated by the use of Theorem A or A^+ at the cost of locally pinching the hyperbolic metric. For a rank two cusp, one can instead first do Dehn surgery of high enough order so that the length of the chosen meridian satisfies Theorem A.

In his tameness proof, in view of Bonahon's theorem, it suffices to assume that the relative core C of $\mathcal{M}(G)$ is a compression body. Extend C to a compact core C' of the PNC manifold M' by adding solid orbifold tori to the cusp tori in the pairing locus and an orbifold 2-handle $\mathbb{S}^2_+ \times [0, 1]$ to the annuli, where $\mathbb{S}^2_+$ denotes a hemisphere. Agol then applied the Orbifold Theorem to C' to get a hyperbolic orbifold structure on M'. Then he used Selberg's Lemma to get a finite cover of M'. If the cover is tame, M' and hence the original $\mathcal{M}(G)$ will be tame as well. See [Bleiler and Hodgson 1996], [Agol 2004] for details.

Now we consider somewhat different situations which also lead, locally, to PNC manifolds. Suppose the compact core of a some $\mathcal{M}(G)$ is a compression body with compressible boundary component S. There is a diskbusting link $\{\alpha_i\} \subset S$. We may assume that $\sum \alpha_i$ is homologous to zero. Denote the mutually disjoint geodesics in $\mathcal{M}(G)$ that are freely homotopic to the elements $\{\alpha_i\}$ by $\{\gamma_i\}$. Unfortunately there is no way of knowing whether or not these geodesics are simple. This problem is eliminated as follows:

Theorem 5.10.9 ("Theorem B" [Canary 1993]). *Fix a small tubular neighborhood U_i about each of the closed geodesics γ_i. The hyperbolic metric of $\mathcal{M}(G)$ can be changed within each U_i so as to obtain a complete PNC manifold M^3 in which each α_i is freely homotopic to a simple geodesic $\gamma_i' \subset U_i$. The PNC metric may be chosen so that in addition its restriction to a thin tubular neighborhood $U_i' \subset U_i$ about γ_i' is hyperbolic.*

As the $\sum \gamma_i'$ of simple geodesics is homologous to zero, an embedded surface $\Sigma \subset M^3$ can be constructed so that $\partial\Sigma = \bigcup \gamma_i$ [Canary 1993, Lemma 3.1]. The surface Σ determines a 2-fold cover $\widehat{M}$ of M^3 branched over $\bigcup \gamma_i$: there is a cyclic group X of order two such that $\widehat{M}/X = M^3$. This is constructed by excising Σ, taking two copies of $\mathcal{M}(G) \setminus \Sigma$ and cross identifying them over Σ. For a simple example consider upper half-space UHS as a 2-sheeted cover of itself under the map $(z, t) \mapsto (z^2, t)$.

By Lemma 5.10.5, $\widehat{M}$ has incompressible boundary! Although the lifted metric in $\widehat{M}$ agrees with the original hyperbolic metric near the ends, it is now singular over the branch lines. After applying Theorem B if necessary to get simple geodesics $\{\gamma_i'\}$, and constructing the cover $\widehat{M}$, the job is completed by removing the branch locus by some more local pinching as follows:

Theorem 5.10.10 ("Theorem C" [Gromov and Thurston 1987; Canary 1993]). *Given the small tubular neighborhood $\cup U_i'^* \subset \widehat{\mathcal{M}}$ about the branch locus of simple geodesics $\bigcup \gamma_i'^*$, $\widehat{M}$ can be given a complete PNC metric with agrees outside $\cup U_i'^*$ with the lifted metric from M^3.*

The bottom line is that by the process outlined, and Lemma 5.10.5, compressible ends can be eliminated at the cost of obtaining a PNC manifold $\widehat{M}$. There remain problems with the topology; see Exercise 5-17.

As already suggested, the tameness of the ends of $\mathcal{M}(G)$ can be dealt with one compressible end at a time — the incompressible ends are already known to be tame (Bonahon's criteria, page 253). For consider a compact core C; each compressible component S of ∂C is the compressible boundary of a compression body C_S, a submanifold of C (Exercise 3-11). Take the covering $\mathcal{M}(H_S)$ of $\mathcal{M}(G)$ determined by the subgroup $\pi_1(C_S)$. The compact core of $\mathcal{M}(H_S)$ can be taken to be the lift of C_S.

Finally we cite an immediate consequence of the Tameness Theorem coupled with the Hyperbolization Theorem (page 324):

Corollary 5.10.11. *If M^3 is a complete PNC manifold of infinite volume, or a noncompact PNC manifold of finite volume, then M^3 has a complete hyperbolic metric.*

For closed PNC manifolds M^3, this is a consequence of Perelman's confirmation of the Geometrization Conjecture (Section 6.4), since M^3 is known to be irreducible, atoroidal, and not Seifert fibered [Cooper and Lackenby 1998].

5-19. *Representation varieties of fuchsian groups.* Fix a fuchsian closed surface group $G = \langle A_1, B_1, \ldots, A_g, B_g\rangle$ with $\prod[A_i, B_i] = 1$. Normalize so that $A_1 = \left(\begin{smallmatrix} k & 0 \\ 0 & k^{-1}\end{smallmatrix}\right)$ and

$B_1 = \binom{a\ c}{c\ d}$, where $k > 1$, $ad - c^2 = 1$, $c \neq 0$. Set

$$R_g = \{\varphi : G \to \varphi(G) \in \mathrm{PSL}(2, \mathbb{R}) \text{ is a normalized homomorphism}\}.$$

Normalized means that $\varphi(A_1), \varphi(B_1)$ are normalized as A_1, B_1. R_g is an irreducible variety of real dimension $6g - 6$. Prove that R_g is a real analytic manifold.

Hint: Consider $H = \varphi(G) \in R_g$. For $2 \leq j \leq g$, choose any X_j, Y_j close to $\varphi(A_j), \varphi(B_j)$. Show that there exist normalized X_1, Y_1 obtained by solving the matrix equation

$$[X_1, Y_1] = Z = \prod_{j=g}^{2} [X_j, Y_j]^{-1}.$$

5-20. *Holomorphic motions.* Suppose $B \subset \mathbb{S}^2$ is an arbitrary set containing at least three points. Let $\{f_\lambda(z) : B \to \mathbb{S}^2\}$ be a family of functions with parameter λ in the open unit disk $\mathbb{D}$. The family is called a *holomorphic motion* of B if the following hold:

- For each fixed $\lambda \in \mathbb{D}$, the map $f_\lambda : z \in B \mapsto f_\lambda(z) \in \mathbb{S}^2$ is one-to-one.
- For each fixed $z \in B$, the map $\lambda \in \mathbb{D} \mapsto f_\lambda(z) \in \mathbb{S}^2$ is holomorphic.
- $f_0(z) = z$ for each $z \in B$.

The great utility of this notion in complex analysis comes from the λ *Lemma* discovered by Ricardo Mané, Paulo Sad, and Dennis Sullivan. Its ultimate expression is due to Slodkowski [1991], to which we refer for history and development; also see [Earle et al. 1994].

The λ-Lemma. *Let G be a Möbius group which maps the set $B \subset \mathbb{S}^2$ onto itself. Suppose $\{f_\lambda\}$ is a holomorphic motion of B. Assume further that, for each $\lambda \in \mathbb{D}$, there is an isomorphism to another Möbius group $\phi_\lambda : G \to G_\lambda$ such that $f_\lambda \circ g(z) = \phi_\lambda(g) \circ f_\lambda(z)$ for all $z \in B$, $g \in G$. Then*

(*i*) *$f_\lambda(z)$ is jointly continuous in λ and z.*
(*ii*) *For fixed λ, $f_\lambda(z)$ is the restriction to B of a K_λ quasiconformal mapping $f_\lambda^* : \mathbb{S}^2 \to \mathbb{S}^2$ which satisfies $f_\lambda^* \circ g(z) = \phi_\lambda(g) \circ f_\lambda^*(z)$ for all $z \in \mathbb{S}^2$, $g \in G$.*
(*iii*) *$\{f_\lambda^*\}$ is a holomorphic motion of $\mathbb{S}^2$.*

Explicitly, $K_\lambda = (1 + |\lambda|)/(1 - |\lambda|)$. Note that continuity in z is not assumed, it is a conclusion. A special case is $G = \mathrm{id}$. On reflection, one finds the conclusions remarkably strong from what at first sight appears as rather weak hypotheses.

Here is the essence of the proof of the λ-Lemma kindly provided by Vlad Markovic. Assume that $f_\lambda(z)$ is an orientation preserving differentiable mapping of a region $B = \Omega$, for each $\lambda \in \mathbb{D}$. Then the complex dilatation $\mu_\lambda(z)$ is defined for all $z \in \Omega$, and it is holomorphic in λ for each fixed z. Moreover $\mu_\lambda(0) = 0$ and $\sup_{z \in \Omega} |\mu_\lambda(z)| \leq 1$. Therefore by the Schwarz Lemma applied as a function of λ, we find that $|\mu_\lambda(z)| \leq |\lambda|$ for all $\lambda \in \mathbb{D}$ and $z \in \Omega$. In particular $f_\lambda(z)$ is quasiconformal in Ω for each $\lambda \in \mathbb{D}$.

An applicable situation might arise as follows. Suppose the generators of the kleinian group G depend analytically on a parameter $\lambda \in \mathbb{D}$ and that there is a corresponding family of isomorphisms $\{\phi_\lambda : G \to G_\lambda\}$ onto other kleinian groups such that $\phi_\lambda(g)$ is parabolic if and only if $g \in G$ is so. For each λ, ϕ_λ then determines an injection f_λ of the set B of loxodromic and parabolic fixed points of G to those of G_λ. We see that $\{f_\lambda\}$ satisfies the hypothesis of the λ-Lemma. Therefore the family $\{G_\lambda\}$ is continuously quasiconformally conjugate to G.

A striking consequence of the λ-lemma is the following. Compare with Corollary 4.1.11.

Theorem 5.10.12 [Sullivan 1985]. *Suppose $h : X \to \mathfrak{R}(G)$ is a holomorphic map of a complex manifold X into the representation variety of a nonelementary kleinian group G such that each $x \in X$ is sent to a group $h(x)$ isomorphic to G. Then for any $x, y \in X$, the groups $h(x), h(y)$ are quasiconformally conjugate.*

That h is holomorphic means that h can be expressed locally as a representation $\rho_x(g)$ which is holomorphic in x for every $g \in G$.

For the quasiconformal deformation space of a geometrically finite group G we ask, is $\mathfrak{T}(G) = \operatorname{Int}\overline{\mathfrak{T}(G)}$? Or is $\mathfrak{T}(G)$ like a ball with a radius removed?

According to [Kapovich 2001, Theorem 8.44], there is a complex manifold V with $\overline{\mathfrak{T}(G)} \subset V \subset \mathfrak{R}(G)$. Apply Theorem 5.10.12 to the inclusion map of the submanifold $\operatorname{Int}\overline{\mathfrak{T}(G)}$ into V. We deduce that all representations in $\operatorname{Int}\overline{\mathfrak{T}(G)}$ are quasiconformal conjugate. Therefore the interior is just $\mathfrak{T}(G)$ itself (see also [McMullen 1998, appendix]).

5-21. Confirm that the Tameness Theorem implies the Ahlfors Finiteness Theorem and the finiteness of the conjugacy classes of parabolics and of finite (elliptic) subgroups as well. To do this make use of the convex core boundary and the fact that there is a sequence of pleated surfaces of uniformly bounded area exiting each geometrically infinite end and relative end; see [Canary 1993, Theorem 8.1].

5-22. *Two generator groups.* Prove by the Tameness Theorem that a two-generator nonelementary kleinian group G (without elliptics) for which $\mathcal{M}(G)$ has infinite volume is a free group (Jaco, Shalen, Agol).

5-23. *Quadratic differentials and measured laminations.* Choose a point $O = (R, \mathrm{id})$ in $\mathrm{Teich}(R)$, where R is a closed Riemann surface of genus $g \geq 2$. A *Teichmüller ray* from O is determined by a holomorphic quadratic differential $\varphi(z)\,dz^2$ on R. The ray consists of the targets of the solutions of the Beltrami equation on R, $F_{\bar{z}} = t(\overline{\varphi(z)}/|\varphi(z)|)F_z$ for $0 \leq t < 1$ with $F(z; 0) = z$ (see Section 2.8). The solution $F_t : R \to R_t$ determines a quadratic differential ψ on R_t. The inverse $F_t^{-1} : R_t \to R$ is a Teichmüller map associated with the quadratic differential $-\psi$ on R_t and $-\varphi$ on R.

A quadratic differential defines a local euclidean metric on R except at its $4g-4$ zeros: $w = \int^z \sqrt{\varphi}\,dz$ is a locally univalent map into $\mathbb{C}$. The preimage of horizontal line segments can be extended globally. We get as a result the *horizontal foliation* Λ_h of R determined by φ. Likewise the inverse images of vertical line segments determine the *vertical foliation* Λ_v. It is customary to normalize quadratic differentials $\{\varphi\}$ so that $\|\varphi\| = \iint_R |\varphi|\, dxdy = 1$.

As we have seen, a Teichmüller map $F : R \to S$ is associated with uniquely determined normalized quadratic differentials φ on R and ψ on S. Each determines a locally euclidean coordinate system, away from its zeros. In terms of these pairs of coordinate systems, F is a locally affine mapping of the form $w = z \mapsto z + t\bar{z}$.

The horizontal foliation Λ_h comes with an associated transverse measure, namely the vertical measure $|dv| = |\mathrm{Im}\, dw|$. Likewise $|du| = |\mathrm{Re}\, dw|$ gives the transverse measure for Λ_v.

For an account of the theory of differentials, see for example [Strebel 1984; Masur 1975; Marden and Strebel 1984; 1986; 1993].

Given a pseudo-Anosov automorphism τ (Exercise 5-6) there is a point $(S, f) \in \mathrm{Teich}(R)$ with the following property. Denote by ψ the differential associated with the Teichmüller map $\tau^* : S \to S$ homotopic to the realization of τ on S. Then $-\psi$ is associated with τ^{*-1}. This means that the Teichmüller geodesic determined by the Beltrami differentials $\{t\bar{\psi}/|\psi| : -1 < t < 1\}$ is the "axis" for τ acting in $\mathrm{Teich}(R)$: τ maps its axis onto itself with attracting fixed point ψ and repelling $-\psi$ in the Thurston boundary, just like the axis of a loxodromic. But we haven't yet associated the quadratic differentials with measured laminations.

The leaves of Λ_h which do not have an endpoint at a zero (the noncritical leaves) have two well determined endpoints when lifted to the universal cover $\mathbb{H}^2$. For each pair of endpoints draw the geodesic with the same end points. Doing this for all noncritical leaves results in a geodesic lamination we again denote by Λ_h. It is equivariant under the group of deck transformations and therefore can also be viewed on R; it has a transverse measure determined by $|dv|$. Likewise the vertical foliation is associated with the measured lamination supported on Λ_v.

We know those measured laminations with supports consisting of simple closed geodesics are dense in $\mathcal{ML}$. What corresponds to closed geodesics are the *simple Jenkins–Strebel differentials*: Given the free homotopy class $[\gamma]$ of a simple closed geodesic $\gamma \subset R$ we can ask, what is the thickest annulus A that can be embedded in the Riemann surface R whose central curves are in $[\gamma]$? In terms of a conformal map of A onto a proper annulus $\{1 < |w| < M\} \subset \mathbb{C}$, the problem is to maximize M over all such embedded A. There exists a unique solution: There is a uniquely determined (normalized) quadratic differential $\varphi[\gamma](z)\, dz^2$ on S such that all of its noncritical horizontal trajectories are simple loops in $[\gamma]$. These horizontal trajectories sweep out an annulus A^*. The complement of A^* is the "critical graph" whose edges are critical trajectories of finite length running between critical points. If A^* is cut along a vertical trajectory segment of $\varphi[\gamma]$, then $w = \int^z \sqrt{\varphi[\gamma]}\, dz$ maps the result onto a rectangle in $\mathbb{C}$. Denote the length of the rectangle by $L[\gamma]$ and the height by $H[\gamma]$; its area is $\|\varphi[\gamma]\| = L[\gamma]H[\gamma] = 1$. The transverse measure to $\Lambda_h = \gamma$ associated with $\varphi[\gamma]$ is $H[\gamma] = 1/L[\gamma]$.

The Teichmüller ray determined by $-\varphi[\gamma]$ consists of surfaces R_t resulting from successively thickening A^* so that in the limit, the result is that R becomes pinched in the class $[\gamma]$.

Extremal length theory in complex analysis implies that for all pairs of simple geodesics on R,

$$L[\alpha] \geq \iota(\alpha, \gamma)H[\gamma], \quad L[\alpha]L[\gamma] \geq \iota(\alpha, \gamma),$$

using the fact that the extremal length of the class $[\alpha]$ is $L[\alpha]^2$ [Strebel 1984].

Now the space of normalized differentials on R is compact, and the simple differentials are dense. From [Kerckhoff 1980] we learn that the functions $L[\gamma]$ extend continuously to all $\mathcal{ML}$, using $L[a\gamma] = aL[\gamma],\ a > 0$.

There exists $c > 0$ such that in comparison with hyperbolic length $\ell(\cdot)$

$$0 < c < \frac{\ell(\gamma)}{L[\gamma]} < \sqrt{4\pi(g-1)} \tag{5.7}$$

for all simple geodesics and hence measured laminations $\gamma \in R$; see [Minsky 1992, §8; Minsky 1993, Lemma 2.1]. The ratio is positive and invariant under scaling by positive constants. As a positive function on the compact set $\mathcal{PML}(R)$, $\frac{\ell(\gamma)}{L[\gamma]}$ has a positive maximum and minimum. The constant on the right comes from an extremal length comparison. Equation (5.7) implies in particular that noncritical horizontal φ-trajectories are quasigeodesics in the hyperbolic metric on S—each lift to $\mathbb{H}^2$ has bounded distance from the hyperbolic geodesic with the same endpoints [Marden and Strebel 1985].

From [Kerckhoff 1980] or [Marden and Strebel 1984, Theorem 5.9] for example we learn that $\lim \varphi[\gamma_n] = \varphi_g$ if and only if $\lim(\gamma_n/c_n) = (\Lambda_g, \mu_g)$ exists. In fact, if we also have $\lim \varphi[\alpha_n] = \varphi_a$ and $\lim(\alpha_n/a_n) = (\Lambda_a, \mu_a)$, then in the limit

$$L_{\varphi_a}(\Lambda_a)L_{\varphi_g}(\Lambda_g) \geq \iota(\mu_a, \mu_g);$$

see [Kerckhoff 1980; Minsky 1993]. Here

$$L_{\varphi_a}(\Lambda_a) = \lim \frac{L[\alpha_n]}{a_n},$$

and correspondingly for $L_{\varphi_g}(\Lambda_g)$.

Moreover, the horizontal laminations corresponding to φ_a and φ_g are Λ_a and Λ_g [Minsky 1994a].

The upshot of these considerations is that a given (projective) measured lamination on R is associated with a uniquely determined normalized quadratic differential. Now suppose we are dealing with an sequence of simple geodesics, or a sequence of pleated surfaces $\{f_n : R \to P_n\}$, exiting an infinite end E of the quasifuchsian manifold based on R. (Or a sequence of hyperbolic metrics $\{\rho_n\}$ converging to a Thurston boundary point.) Represented on R, the sequence of simple geodesics $\{\gamma_n\}$, or sequence of bending laminations, is converging to the ending lamination (Λ_E, μ). Correspondingly the normalized quadratic differentials converge $\lim \varphi[\gamma_m] = \varphi[\Lambda_E]$. As seen below, we do not need to pass to subsequences.

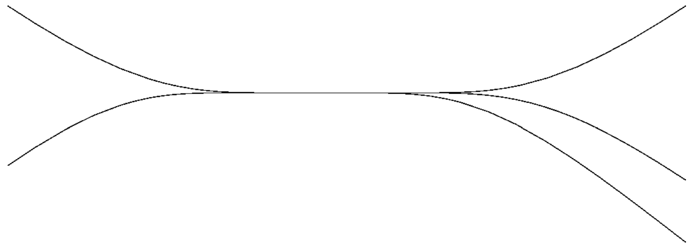


Fig. 5.18. The local structure of a train track.

Under the assumption of bounded geometry and incompressible ends, Minsky [1993; 1994a] proved the following. Suppose we have a sequence of pleated surfaces $f_n : R \to P_n$ exiting E. The hyperbolic structures on $\{P_m\}$ are of bounded distance from the Teichmüller rays determined by $\{-\varphi[\gamma_m]\}$. The horizontal trajectories of $\varphi[\Lambda_E]$ is equivalent to the ending lamination Λ_E. By way of analogy, if $\gamma \subset R$ is a simple loop then the Teichmüller ray determined by $-\varphi[\gamma]$ thickens the annular region about γ so that the corresponding sequence of Riemann surfaces R_t pinches along γ. In the quasifuchsian manifold, the corresponding ending lamination is just the geodesic represented by $\gamma \in R$.

The Teichmüller ray determined by $-\varphi[\Lambda_E]$ in $\mathrm{Teich}(R)$ on the other hand projects to a compact subset of moduli space $\mathrm{Teich}(R)/\mathcal{M}(R)$, because of the bounded injectivity radius hypothesis. This implies that Λ_E is uniquely ergodic, according to [Masur 1992, Theorem 1.1]. Therefore the differential $\varphi[\Lambda_E]$ is uniquely determined by Λ_E.

Thus the sequence of pleated surfaces are being "pinched" to the pleating loci — the horizontal trajectories of $\varphi[\gamma_n]$. Along the sequence, the measure of a given transverse segment is increasing without bound. For more discussion see Section 6.1.1.

We have touched on the "dictionary" between measured foliations in topology, measured laminations in geometry, quadratic differentials in complex analysis, and there are also train tracks in combinatorics, as we will see next.

5-24. *Train tracks.* Suppose S is a closed hyperbolic surface. A *train track* $\tau \in S$ is a finite 1-dimensional graph such that all vertices are trivalent. The relation of vertices to edges is to be like a switching point for a train. The three edges e_1, e_2, e_3 at a vertex v are placed so that a train coming in on either track e_1 or e_2 must exit on e_3, and conversely, a train coming in on track e_3 can exit on either e_1 or e_2.

Formally, the edges are C^1-arcs and the tangent lines have one sided limits at their end points. At each vertex, the tangent lines of the three edges coincide (thus there is one line ℓ_v at each vertex so that ℓ is the limit of the tangent lines to all three tracks at v).

It is also assumed that each component of $S \setminus \tau$ is a triangle.

A train track with *weights* has numbers $c > 0$ assigned to each edge. The *switch condition* is that at a vertex v, the numbers c_1, c_2, c_3 assigned to the three edges must satisfy $c_1 + c_2 = c_3$, using the labeling introduced above. If these are integers we can interpret the assumption to be that e_1 and e_2 each carry c_1 and c_2 parallel tracks coming in to v, and e_3 carries $c_1 + c_2$ parallel tracks leaving v. Thus if all the assigned numbers are integers, a particular train can take many possible journeys over the set of tracks. The journey will be of finite length before the trip repeats itself. If the weights are not integers, the journey of a train may be infinitely long.

A train track with weights uniquely determines a measured foliation on S: the leaves run along the branches with transverse measure given by the weights. Conversely, by pinching together nearly parallel leaves, every measured foliation is represented by some track τ. More precisely, a foliation F is mapped onto τ if there is a map ϕ of $S \setminus \{\text{singularities of } F\}$ such that ϕ is homotopic to the identity in such a way that tangent lines to leaves of F are sent to tangent lines of τ.

Likewise each measured geodesic lamination can be represented by a weighted train track, and conversely, each weighted train track uniquely determines a measured lamination.

The theory of train tracks was created by Thurston [1979, §§8.9, 9.7]. For an extended exposition of the theory see [Penner and Harer 1992].

5-25. *Extension of boundary deformations to* $\mathcal{M}(G)$. The purpose of this exercise is to sketch the proof of Theorem 5.1.3. We have to explain the relation of the quasiconformal deformation space $\mathfrak{T}(G)$ to the product of the classical Teichmüller spaces of the components $\{S_i\}$ of $\partial\mathcal{M}(G)$: $\mathrm{Teich}(S_1) \times \cdots \times \mathrm{Teich}(S_k)$.

We recall that two normalized quasiconformal deformations of G are equivalent ($F_1 \sim F_2$) if they induce the same isomorphism $\varphi : G \to H$. This means in terms of their projections $f_1, f_2 : \partial\mathcal{M}(G) \to \mathcal{M}(H)$, that $f_2^{-1} \circ f_1$ extends to $\mathcal{M}(G)$ and is homotopic to the identity on $\mathrm{Int}\,\mathcal{M}(G)$; see Section 3.7.2 and Exercise 5-16.

We also have to consider the stronger equivalence, namely

$$F_1 \simeq F_2 \;\Leftrightarrow\; f_2^{-1} \circ f_1 : \partial\mathcal{M}(G) \to \partial\mathcal{M}(G) \text{ is homotopic to } \mathrm{id}\,.$$

It follows that $F_1 \sim F_2$ in the earlier definition. These two equivalences differ only when $\partial\mathcal{M}(G)$ is compressible.

To mirror the difference in the two equivalence relations we introduce the group $X(G)$ consisting of normalized quasiconformal deformations that preserve each component of $\Omega(G)$ and induce the identity automorphism of G. Here we refer to Theorem 3.7.3.

From the point of view of the manifolds, $X(G)$ consists of equivalence classes of quasiconformal automorphisms $h : \partial\mathcal{M}(G) \to \partial\mathcal{M}(G)$ that extend to $\mathcal{M}(G) \to \mathcal{M}(G)$ and which are homotopic to the identity on the interior $\mathrm{Int}\,\mathcal{M}(G)$. Two such maps h_1, h_2 are to be identified if and only if $h_2^{-1} \circ h_1$ is homotopic to the identity on $\partial\mathcal{M}(G)$ too; specifically, $h_2^{-1} \circ h_2$ maps each component S_i onto itself and is homotopic on S_i to the identity.

If $\mathcal{M}(G)$ is boundary incompressible, $X(G) = \mathrm{id}$.

Denote by $\mathrm{Mod}_0(S_i)$ the group of homotopy classes of quasiconformal mappings $h : S_i \to S_i$ which extend to $\mathcal{M}(G)$ to be homotopic in $\mathrm{Int}\,\mathcal{M}(G)$ to the identity — Theorem 3.7.3 again. To be more precise, such a map h in particular fixes the punctures on S_i, and the set of compressing loops. Extend h from S_i to all $\partial\mathcal{M}(G)$ by setting it equal to the identity on S_m, $m \neq i$. Then h extends to $\mathcal{M}(G)$ and is homotopic in $\mathrm{Int}\,\mathcal{M}(G)$ to the identity. In other terms, h is the projection of a quasiconformal automorphism h^* of each component $\Omega_{i,j}$ over S_i with the property that h^* induces the identity automorphism of $\mathrm{Stab}(\Omega_{i,j})$ and extends continuously to the identity map of $\partial\Omega_{i,j}$. The group $\mathrm{Mod}_0(S_i)$ is a subgroup of the mapping class group $\mathfrak{M}(S_i)$.

Therefore the group $X(G)$ splits into a direct product

$$X(G) = \mathrm{Mod}_0(S_1) \times \mathrm{Mod}_0(S_2) \times \cdots \times \mathrm{Mod}_0(S_k).$$

The group $\mathrm{Mod}_0(S_i)$ acts without fixed points on $\mathrm{Teich}(S_i)$. For suppose, for example, that $h \in \mathrm{Mod}_0(S_i)$ fixes the origin (S_i, id) in $\mathrm{Teich}(S_i)$. Then h is homotopic to a conformal map $h_0 : S_i \to S_i$. Now h and then h_0 lift to automorphisms h^* and h_0^* of $\Omega_{i,j}$ over R_i; we can choose h^* to be homotopic in $\Omega_{i,j}$ to h_0^*. We know that h^* extends continuously to $\partial\Omega_{i,j}$ and fixes every point, Exercise 3-34. So the same is true of h_0^* which therefore must be the identity since it is a conformal automorphism. Consequently h^* is homotopic in $\Omega_{i,j}$ to the identity and h is homotopic in S_i to the identity.

The classical results obtained by projection from the space of Beltrami differentials with respect to G on $\Omega(G)$ that imply $\mathfrak{T}(G)$ is a complex analytic manifold [Ahlfors 1966].

Examine now the quotient $\mathrm{Teich}(S_i)/\mathrm{Mod}_0(S_i)$. Here we are identifying those elements of the Teichmüller space of S_i that are related by a mapping that is the identity with respect to the interior of the 3-manifold.

Since we are taking the quotient of an analytic manifold by a discrete group of fixed point free biholomorphic automorphisms, $\mathrm{Teich}(S_i)/\mathrm{M}_0(S_i)$ is an analytic manifold of the same dimension as $\mathrm{Teich}(S_i)$.

This completes the proof of Theorem 5.1.3.

Remark 5.10.13. Suppose all components $\Omega_{i,j}$ of $\Omega(G)$ are simply connected but that there may be torsion in their stabilizers $G_{i,j}$. Then in addition to the punctures on each component $S_i \subset \partial\mathcal{M}(G)$ there will be $b_i \geq 0$ cone points. In this case the dimension count will be

$$\sum (3g_i + b_i + n_i - 3).$$

For it is an interesting fact that $\mathrm{Teich}(S_i)$ is biholomorphically equivalent to $\mathrm{Teich}(S_i')$ where S_i' is the result of removing the cone points. That is, the dimension is the same whether you have $b_i + n_i$ punctures, or n_i punctures and b_i cone points [Marden 1969], [Bers and Greenberg 1971].

The basis for the equivalence is the following fact: A homeomorphism $f : S_i \to S_i$ lifts to a homeomorphism f^* of $\Omega_{i,j}$ which induces the identity automorphism of

$G_{i,j}$ if and only if f is homotopic in S'_i to the identity map. This follows from the fact that γ is freely homotopic in S'_i to $f(\gamma)$ for all simple loops $\gamma \subset S'_i$.

Consequently Theorem 5.1.3 remains true at least when the components of $\Omega(G)$ are simply connected but $\{G_{i,j}\}$ contains elliptics and the deformations preserve elliptics and their orders. The original papers [Bers 1970b; Maskit 1971; Kra 1972] include the general case.

6

Hyperbolization

In this chapter we will explain the Hyperbolization Theorem for 3-manifolds, one of the truly great mathematical discoveries of the twentieth century. This theorem shows that the interiors of most compact 3-manifolds can be realized as kleinian manifolds. As a consequence, such 3-manifolds can be described and classified not just in terms of their topology, but more powerfully, in terms of their geometrical properties — their shape.

6.1 Hyperbolic manifolds that fiber over a circle

6.1.1 Automorphisms of surfaces

We begin by reviewing some facts about automorphisms of surfaces. We will continue using as basepoint a fuchsian group G and associated Riemann surfaces $R = \mathrm{LHP}/G$, $R' = \mathrm{UHP}/G$, closed with at most a finite number of punctures. Suppose $\alpha : R \to R$ is an orientation preserving automorphism which is not homotopic to the identity. As we learned in §5.5.1, the automorphism α, or rather its homotopy (and isotopy) class, induces an automorphism of the Bers slice $\mathcal{B}(R)$ based on R, which we also denote by α, by the action

$$\alpha : (S_{\mathrm{bot}} = R,\, S^{\mathrm{top}};\, J) \mapsto (S_{\mathrm{bot}} = R,\, S^{\mathrm{top}};\, J \circ \alpha).$$

The map α does not change the conformal type of either the bottom or the top Riemann surface. Instead the *relationship* between the two surfaces as dictated by $J \circ \alpha$ changes (since α is not homotopic to the identity). The group of homotopy classes of orientation preserving automorphisms α is called the *mapping class group* or *Teichmüller modular group*. It is the group of all isometries of Teichmüller space in the Teichmüller metric of §2.8.

If instead R is a torus it has a continuous group of automorphisms. After quotienting out this group, in effect fixing a point x on the torus, the Teichmüller space $\mathrm{Teich}(R)$ of a torus can be identified with the upper half-plane and the corresponding Teichmüller modular group becomes the classical modular group (see Exercises 2-6 and 5-4). In many respects the general mapping class group is analogous to this one.

Recall from Exercise 5-6 that a *pseudo-Anosov mapping* is a homeomorphism α of a surface R onto itself with these properties: (i) No power α^n is homotopic to the identity; and (ii) α does not preserve the set of free homotopy classes of any system of mutually disjoint, simple loops on R (none of which is homotopic to a point or to a puncture). Such automorphisms of R are the "generic" automorphisms. For more information see [Thurston 1988; Fathi et al. 1979] and Exercise 5-6.

The automorphism α acts on simple loops and, by passing to the closure, on the space of projective measured laminations $\mathcal{PML}(R)$. It has exactly two fixed points. Namely for any simple geodesic γ and hyperbolic length $\ell(\cdot)$ on R,

$$\mu_{\text{attr}} = \lim_{n\to+\infty} \frac{\alpha^n(\gamma)}{\ell(\alpha^n(\gamma))}, \quad \mu_{\text{rep}} = \lim_{n\to-\infty} \frac{\alpha^n(\gamma)}{\ell(\alpha^n(\gamma))}.$$

This situation is seen as analogous to Anosov maps of the torus (Exercise 5-4), especially as expressed in the context of the theory of quadratic differentials (Exercise 5-23).

From the point of view of $\mathcal{B}(R) \cup \partial_{\text{th}}$, a pseudo-Anosov α has a unique attracting and repelling fixed point, $(\Lambda_{\text{attr}}, \mu_{\text{attr}})$, $(\Lambda_{\text{rep}}, \mu_{\text{rep}})$ each of which lies on the Thurston boundary ∂_{th}. There is a unique "axis" of α in $\mathcal{B}(R)$: a geodesic in the Teichmüller metric whose endpoints are the fixed points (see Exercise 5-23). The axis is left invariant by the action of α. Indeed for any point $P \in \mathcal{B}(R) \equiv \text{Teich}(R)$, $\lim_{n\to+\infty} \alpha^n(P) = \mu_{\text{attr}}$ and $\lim_{n\to-\infty} \alpha^n(P) = \mu_{\text{rep}}$.

In the notation of page 280, start with the hyperbolic structure ρ_0 on R. The iterates $\{\alpha^n\}$ determine a sequence of new hyperbolic structures $\{\rho_n\}$ on R, namely α^n sends the point $P = (R, \text{id}) \in \text{Teich}(R)$ to $P_n = (R, \alpha^n)$. Let γ_n denote the geodesic on R freely homotopic to $\alpha^n(\gamma)$. In the respective lengths, $\ell_{\rho_n}(\gamma_n) = \ell_{\rho_0}(\gamma)$ while $\rho_0(\gamma_n) \to \infty$. Furthermore as $n \to +\infty$, $\{\gamma_n/\ell_{\rho_0}(\gamma_n)\}$ converges in $\mathcal{PML}(R)$ to $(\Lambda_{\text{attr}}, \mu_{\text{attr}})$, and as $n \to -\infty$ to $(\Lambda_{\text{rep}}, \mu_{\text{rep}})$ [Otal 1996, §1.5].

These projective measured laminations are the *stable* and *unstable* measured laminations for α. By analogy, for the affine map $A : x \mapsto K^{1/2}x,\ y \mapsto K^{-1/2}y$, $K > 1$, the x-axis is the stable lamination and the y-axis is the unstable — for most points $p \in \mathbb{R}^2$, $\lim_{m\to+\infty} A^m(p)$ lies on the x-axis. Also note that the length of a transverse segment to the x-axis is decreased by the factor $K^{-1/2}$ while the length of a transverse segment to the y-axis is increased by $K^{1/2}$. This phenomenon equally true for the transverse measures to the attracting and repelling (stable and unstable) fixed points of α. In fact if α is a pseudo-Anosov acting on $\text{Teich}(R) \cup \partial_{\text{th}}$ with fixed points $\mu_{\text{attr}}, \mu_{\text{rep}}$, there exists $K > 1$ such that for any simple closed geodesic γ, the generalized intersection numbers satisfy

$$\iota(\alpha(\gamma), \mu_{\text{attr}}) = K^{-1}\iota(\gamma, \mu_{\text{attr}}) \quad \text{and} \quad \iota(\alpha(\gamma), \mu_{\text{rep}}) = K\,\iota(\gamma, \mu_{\text{rep}});$$

see [Otal 1996, §1.5].

From these relations, one can draw the expected conclusions that any nonzero power of a pseudo-Anosov is a pseudo-Anosov, and the homotopy class of no loop γ is fixed by a pseudo-Anosov. In analogy to the case of loxodromic Möbius transformations, (i) the fixed points of two pseudo-Anosovs are either distinct or identical, and (ii) if α_1, α_2 have distinct fixed points on ∂_{th}, then $\langle \alpha_1^m, \alpha_2^n \rangle$ is a free abelian group for sufficiently large $m, n > 0$ [Ivanov 1992].

The sequence of groups in $\mathcal{B}(G)$ corresponding to the triples $\{(R, R'; J \circ \alpha^n)\}$ converges algebraically to a singly degenerate group $H_{\mathrm{attr}} \in \partial\mathcal{B}(R)$ as $n \to -\infty$. This is because the action of $J \circ \alpha^n$ on the top surface R' is homotopic to the action of α^{-n} directly on R', since J is orientation reversing. The limit Λ_{attr} of the sequence $\{\alpha^{-n}(R')\}$ is the ending lamination of the top (geometrically infinite) end of $\mathcal{M}(H_{\mathrm{attr}})$. There are no new parabolics in H_{attr} so the convergence is not only algebraic but is also geometric by Theorem 4.6.2.

If instead $n \to +\infty$, the corresponding points of $\mathcal{B}(R)$ converge algebraically to a different singly degenerate group $H_{\mathrm{rep}} \in \partial\mathcal{B}(R)$ with ending lamination Λ_{rep}. The two laminations $\Lambda_{\mathrm{attr}}, \Lambda_{\mathrm{rep}}$ *fill up* the reference surface R: they have no leaves in common and each complementary component $R \setminus \Lambda_{\mathrm{attr}} \cup \Lambda_{\mathrm{rep}}$ is a polygon possibly containing a single puncture. An alternate description is that $\iota(\nu, \mu_{\mathrm{attr}}) + \iota(\nu, \mu_{\mathrm{rep}}) > 0$ for any measured lamination $(\Lambda, \nu) \neq 0$.

Return now to the full quasifuchsian deformation space $\mathfrak{T}(G)$. Consider the sequence of quasifuchsian groups given by $(\alpha^m(R), R'; J \circ \alpha^{-m-n})$, $m, n \to +\infty$. The top surface R' is related to R as before since on R, $\alpha^{m-m} = \mathrm{id}$, but independently we are applying α^m to R. As a consequence of the Double Limit Theorem to follow, the sequence of groups converges algebraically (and also geometrically) to a doubly degenerate group $H \in \partial\mathfrak{T}(G)$. The ending laminations for the top and bottom ends of $\mathcal{M}(H)$ are $\Lambda_{\mathrm{attr}}, \Lambda_{\mathrm{rep}}$ [Thurston 1986c, §4; McMullen 1996, §§3.3-5]. We are applying the Ending Lamination Theorem , which obviates the necessity of taking subsequences.

6.1.2 The Double Limit Theorem

Double Limit Theorem [Thurston 1986c]. *Let* $(\Lambda_{\mathrm{bot}}, \mu_{\mathrm{bot}})$ *and* $(\Lambda_{\mathrm{top}}, \mu_{\mathrm{top}})$ *be points of* $\partial_{\mathrm{th}}\mathrm{Teich}(R)$, *where* Λ_{bot} *and* Λ_{top} *fill* R. *Suppose they are, respectively, the limits of the sequences of hyperbolic structures* $\{p_i\}$ *and* $\{q_i\}$ *in* $\mathrm{Teich}(R)$. *Let* H_i *be a normalized quasifuchsian group whose bottom surface carries* p_i *and top surface* q_i *where the natural involution* J *between the top and the bottom interchanges the markings on* p_i *and* q_i. *Then* $\{H_i\}$ *converges algebraically to a group* $H \in \partial\mathfrak{T}(G)$. *The ending laminations of* $\mathcal{M}(H)$ *are* Λ_{bot} *and* Λ_{top}, *respectively.*

In fact from §5.9.4, there are sequences $\{\mu_m, \nu_n\}$ of measured laminations converging to $\mu_{\mathrm{bot}}, \mu_{\mathrm{top}}$ such that

$$\lim_{m\to\infty} \mathrm{Len}_{p_m}(\mu_m) = \lim_{n\to\infty} \mathrm{Len}_{q_n}(\nu_n) = 0.$$

This reinforces the picture of "pinching" the approximating surfaces along the ending laminations.

Convergence to a point on the Thurston boundary $\partial_{\mathrm{th}}\mathrm{Teich}(R)$ is discussed in §5.9.4. If R has punctures the laminations are of course confined to a compact submanifold. We have modified the original statement by bringing in the Ending Lamination Theorem , which releases us from the obligation of passing to subsequences. The limit group H will be doubly degenerate without new parabolics only if both laminations are arational. If instead Λ_{bot}, Λ_{top} are transverse pants decompositions, the limit group H is a maximal cusp; the two boundary components correspond to the result of pinching the top and the bottom along the respective pants loops.

Doubly degenerate groups appear as subgroups of hyperbolic 3-manifolds that fiber over the circle, as we will see below. The first doubly degenerate group appears in [Jørgensen 1977a], it was given by explicit generating matrices. For other explicitly constructed degenerate groups see [Jørgensen and Marden 1979]. For a generalization to compression bodies (function groups) see [Kleineidam and Souto 2002].

6.1.3 Manifolds fibered over the circle

Suppose now G is a fuchsian or quasifuchsian group so that $\mathcal{M}(G) \cong R \times [0, 1]$. Let R, R' denote the bottom and top components of $\partial\mathcal{M}(G)$. Suppose $\tau : R \to R'$ is an orientation reversing map. The 3-manifold M^3 without boundary that results from identifying the boundary components via τ, namely $M^3 = \mathcal{M}(G)/\sim\tau$, is fibered over a circle. The fibers are surfaces homeomorphic to R; the "circle" is the image of a simple arc in $\mathcal{M}(G)$ connecting a point $x \in R$ to $\tau(x) \in R'$. The map τ factors as $\tau = J \circ \alpha : R \to R'$. Here J is the orientation reversing involution that exchanges R and R' and induces the identity on $\pi_1(\mathcal{M}(G))$ while α is an automorphism of R.

For our purposes, a better way of describing M^3 is as a mapping torus. Namely let $\alpha : R \to R$ be a homeomorphism and form the *mapping torus*

$$M^3 = R \times \mathbb{R}/\langle (x, t) \mapsto (\alpha(x), t + 1)\rangle.$$

In the universal cover $R \times \mathbb{R}$, α determines an infinite cyclic group of deck transformations which are translations taking one lift of R to another.

The first case of the hyperbolization theorem has the following beautifully succinct statement.

Manifolds Fibered over the Circle [Thurston 1986c; Otal 1996; Kapovich 2001]. *Necessary and sufficient for the manifold $\mathcal{M}(G)/\sim\tau$ to have a hyperbolic structure $\mathcal{M}(X)$ is that $\alpha : R \to R$ be pseudo-Anosov. In this case, $\mathcal{M}(X)$ has finite volume.*

In contrast, consider what happens if α is not pseudo-Anosov. Suppose it fixes the free homotopy class of a simple geodesic $c \in R$. Then c and $\tau(c)$ bound a cylinder C in $\mathcal{M}(G)$. The cylinder C rolls up to form an incompressible torus in M^3. This is possible in a hyperbolic manifold only if the torus comes from a rank two parabolic, which is the case only when c encircles a puncture.

The lifts of the fibers to $\mathbb{H}^3$ are fascinating. Suppose $\mathcal{M}(X)$ is fibered over a circle with fibers homeomorphic to a finitely punctured, closed surface R. Choose a fiber Y. The lifts — components of the preimage — $\{Y^*\}$ of Y to $\mathbb{H}^3$ form a discrete set of mutually disjoint simply connected surfaces. There is a Möbius transformation $T \in X$

such that if Y^* is one lift, the orbit of Y^* under $\langle T \rangle$ comprises the complete set. In fact $X = \langle H, T \rangle$ and T represents the pseudo-Anosov α that determines $\mathcal{M}(X)$. Set $H = \mathrm{Stab}_X(Y^*)$; then $THT^{-1} = H$.

We claim that $\Omega(H) = \varnothing$. Otherwise, choose $K \subset \Omega(H)$ compact. Then $K \cap \overline{T^n(Y^*)} = \varnothing$ for all n. In particular, no fixed point of T lies in K. Therefore $K \subset \Omega(X)$, which is impossible. Consequently H is a periodic doubly degenerate group without parabolics, isomorphic to the fuchsian model G.

The Möbius transformation T projects to an automorphism Φ of $\mathcal{M}(H)$. The sequence of planes $\{T^n(Y^*)\} \subset \mathbb{H}^3$ projects to a discrete Φ-invariant sequence of surfaces $\{Y_n\} \subset \mathcal{M}(H)$ exiting its two ends. If $\gamma \subset Y$ is a simple loop, and γ_g is its geodesic representative in $\mathcal{M}(H)$, then the two ending laminations are determined by the exiting sequences of equal length geodesics $\lim_{n\to+\infty} \Phi^n(\gamma_g)$ and $\lim_{n\to-\infty} \Phi^n(\gamma_g)$; see [Minsky 2003b]. These are the "fixed points" of the pseudo-Anosov α, as described earlier.

We note that $\mathcal{M}(H)$ has the property of bounded geometry, see Section 5.6.2.

It is shown in [Cannon and Thurston 1989] (see also [Minsky 1994a; Mitra 1998a; 1998b]) that there exists a quasiisometric map $f : \mathbb{H}^2 \to Y^*$, which induces an isomorphism $G \to H$ and extends continuously to a map $\mathbb{S}^1 \to \partial Y^* \subset \mathbb{S}^2$, which is therefore a space-filling (Peano) curve. It is the image of a collapsing map of $\partial\mathbb{H}^2$ with respect to the two laminations associated with the pseudo-Anosov — placing one in the upper half-plane, for example, and the other in the lower.

It was originally believed that manifolds fibered over the circle could not be hyperbolic because of the strange properties their coverings would have. Thus Jørgensen's example of a periodic doubly degenerate groups with fiber the once-punctured torus was instrumental in inspiring the early development of the subject (see [Thurston 1986c, §0]). An oft cited, closely related example is the hyperbolic manifold, also fibered over the circle with once-punctured torus fibers, which is homeomorphic to the complement of the figure-8 knot (p. 164). See also [Jørgensen and Marden 1979].

So, starting with the fuchsian G and pseudo-Anosov α, to find H and T we have to move through the deformation space $\mathfrak{T}(G)$ until we find H on its boundary with ending laminations associated with α. H will be "periodic" with respect to a loxodromic T representing α.

Thurston asked whether every hyperbolic manifold of finite volume is virtually fibered; that is, whether each has a finite cover that is fibered over the circle. A necessary condition that a manifold be fibered over the circle is that it has infinite homology, which is automatically satisfied for cusped manifolds. In [Button 2005] one finds a list of more than 100 closed manifolds that are themselves not fibered but which have finite covers which are fibered. This is done by finding fibered manifolds which are commensurable with nonfibered ones. One of those found has infinite homology. Earlier it was discovered that over 87% the manifolds in the Callahan–Hildebrand–Weeks census [Hildebrand and Weeks 1989] of nearly 5000 orientable cusped finite volume manifolds are themselves fibered. The question remains wide open.

6.2 The Skinning Lemma

The key to finding hyperbolic structures on a large class of 3-manifolds lies in being able to find fixed points of certain mappings of deformation spaces $\mathfrak{T}(G)$ onto themselves. This method does not work in finding hyperbolic structures for manifolds that fiber over a circle, which is why they are treated separately. In this section we will give an exposition of the procedure for finding the needed fixed points.

6.2.1 Hyperbolic manifolds with totally geodesic boundary

Suppose $\mathcal{M}(G)$ is a geometrically finite, acylindrical (and hence boundary incompressible) manifold. Because of the hypothesis, if Ω is a component of $\Omega(G)$, the subgroup $G_\Omega = \mathrm{Stab}(\Omega)$ is a quasifuchsian group — see Exercise 3-10. Now the ordinary set of G_Ω has two components one of which is Ω. Denote the other by $\Omega' = \mathbb{S}^2 \setminus \overline{\Omega}$. Since G itself cannot be quasifuchsian, Ω' is not a component of $\Omega(G)$. There is an orientation reversing quasiconformal involution $\sigma : \Omega' \leftrightarrow \Omega$ that induces the identity automorphism of G_Ω. Its extension pointwise fixes the common boundary. Its projection is an orientation reversing map $\sigma : S = \Omega/G_\Omega \leftrightarrow S' = \Omega'/G_\Omega$.

Choose a fundamental set of components $\Omega_1, \dots, \Omega_r$ of $\Omega(G)$ in the sense that no two are equivalent under G yet their G-orbits cover $\Omega(G)$. Label the corresponding Riemann surfaces $S_1, \dots, S_r$; these are the boundary components of $\mathcal{M}(G)$. We have an associated set of Riemann surfaces and corresponding orientation reversing quasiconformal involutions which we will write as

$$\sigma : (S_1, \dots, S_r) \mapsto (S'_1, \dots, S'_r).$$

The map σ is called the *skinning map* since it removes the "skin" that hides the structures $\{S'_i\}$ below, as skinning an apple exposes the yummy stuff underneath.

Next, define a map ρ that operates on r-tuples of Riemann surfaces by the following operation:

$$\rho : (R_1, \dots, R_r) \mapsto (\bar{R}_1, \dots, \bar{R}_r),$$

Here $\bar{R}_i$ denotes the Riemann surface obtained from R_i by replacing each local coordinate $\{z\}$ by its complex conjugate $\{\bar{z}\}$. To be more concrete, suppose Γ_i is a fuchsian group acting in LHP and UHP such that LHP/$\Gamma_i = R_i$. Then $\bar{R}_i = \mathrm{UHP}/\,\Gamma_i$.

The question is: Given $\mathcal{M}(G)$ and its boundary components $S_1, \dots, S_r$ consider the orientation preserving map

$$\rho \circ \sigma : (S_1, \dots, S_r) \mapsto (\bar{S}'_1, \dots, \bar{S}'_r).$$

We can look at this as a homeomorphism of $\mathrm{Teich}(S_1) \times \cdots \times \mathrm{Teich}(S_r)$ onto itself. Can we deform $\mathcal{M}(G)$ by a quasiconformal deformation to get a manifold $\mathcal{M}(G^*)$ for which $\rho \circ \sigma$ is the identity? If so, at the fixed point, $(\bar{S}'_1, \dots, \bar{S}'_r) \equiv (S_1, \dots, S_r)$.

For simplicity, assume that $r = 1$. Consider the situation up in $\Omega(G^*)$. Choose a component Ω; all other components are conjugate by an element of G^*. We have the situation,

$$\Omega \xrightarrow{\sigma} \mathbb{S}^2 \setminus \overline{\Omega} \xrightarrow{\rho} \Omega.$$

Now ρ is an orientation reversing, angle preserving map that commutes with Stab(Ω) and pointwise fixes $\partial\Omega$. Necessarily $\partial\Omega$ is a round circle and ρ is the reflection in $\partial\Omega$.

The term used is that the boundary $\partial\mathcal{M}(G^*)$ is *totally geodesic*. This is a small abuse of terminology however. What is really meant is that the convex core of $\mathcal{M}(G^*)$ is bounded by totally geodesic surfaces, each the dome over some S_i. A totally geodesic surface S^* embedded in the interior of $\mathcal{M}(G^*)$ is one with the property that the geodesic between any two points on S^* lies within S^*. This occurs if and only if each lift of S^* in $\mathbb{H}^3$ is a hyperbolic plane. Cool!

The geodesic boundary of $\mathcal{M}(G^*)$ makes it possible to directly construct its hyperbolic double. Look at a typical component Ω of $\Omega(G^*)$. Let J denote the reflection in the bounding circle $\partial\Omega$. J is also the reflection of $\mathbb{H}^3$ in Dome(Ω). Consider the new group $\langle G, JGJ\rangle$. J conjugates the action of G in the exterior of Ω and the "outside" of its dome to new action in Ω and "under" the dome. But also $JgJ = g$ for each element $g \in$ Stab(Ω). Topologically we have attached two copies of $\mathcal{M}(G^*)$ along the common boundary $S = \Omega/\mathrm{Stab}(\Omega)$. When this is done for all the boundary components we will have built a kleinian group representing the double, which is a manifold of finite volume. Finding a manifold with totally geodesic boundary is equivalent to finding one whose double has a hyperbolic structure — can be represented by a kleinian group.

Theorem 6.2.1 [Thurston 1982a; McMullen 1990]. *If $\mathcal{M}(G)$ is geometrically finite and acylindrical, there is a unique manifold $\mathcal{M}(G^*)$ in its quasiconformal deformation space $\mathfrak{T}(G)$ which has totally geodesic boundary.*

Without the acylindrical requirement the theorem would be false. For example, in a quasifuchsian space there are many fuchsian deformations of a fuchsian group.

It is because of Mostow's Rigidity Theorem applied to the double, that $\mathcal{M}(G^*)$ is unique. The first proof by Thurston [1982a] applied the Hyperbolization Theorem (see page 324 below) to the topological double of $\mathcal{M}(G)$. See also Exercise 6-2.

Recently Peter Storm answered a related conjecture of Bonahon about convex core volumes when he proved:

Theorem 6.2.2 [Storm 2002b; 2002a]. *Suppose $\mathcal{M}(G)$ is geometrically finite, acylindrical and without rank two cusps. In the quasiconformal deformation space of $\mathcal{M}(G)$, the volume of the convex core $\mathcal{C}(G)$ is uniquely minimal for the manifold $\mathcal{M}(G^*)$ with totally geodesic boundary.*

An analogous result for surfaces was discussed in Exercise 4-17.

David Wright's Figure 6.2.1 is a wonderful illustration of this section. The limit set is a *Sierpiński gasket*, that is, it is a closed subset of $\mathbb{S}^2$ with empty interior whose complement is the union of round disks with mutually disjoint closures. The complementary set is called a *Sierpiński carpet*. The group G has no parabolics and is constructed by identifying the faces of two adjacent truncated ideal tetrahedra. See [Thurston 1997, p. 133] or http://www.math.okstate.edu/~wrightd/Marden for details. It is known that there are exactly eight nonisometric manifolds with genus two geodesic boundaries formed from two ideal tetrahedra [Fujii 1990]. Their convex cores all have the same volume. More generally, there are 151 manifolds with geodesic boundary constructed from three tetrahedra, and 5,033 ones constructed from four [Frigerio et al. 2004].

6.2.2 Skinning the manifold (Part II)

We need a theorem for a more general situation. The most general case is of a finite collection of geometrically finite $\{\mathcal{M}(G_i)\}$. Suppose τ is an orientation reversing quasiconformal involution of $\bigcup \partial\mathcal{M}(G_i)$ that has the effect of preserving the set of solid pairing tubes. $\bigcup P_i$. Form the new 3-manifold $M_\tau = \bigcup \mathcal{M}(G_i)/\sim \tau$. Assume that M_τ is connected. In this generality, the case leading to totally geodesic boundaries is included.

However for simplicity assume that $\tau : \partial\mathcal{M}(G) \to \partial\mathcal{M}(G)$ is an orientation reversing quasiconformal involution (sending punctures to punctures) that sends S_i to S_j and S_j to S_i where $j = j(i) \neq i$.

Form the topological manifold $M_\tau = \mathcal{M}(G)/\sim \tau$ by gluing the boundary components as prescribed by τ. We sail through the quasiconformal deformation space $\mathfrak{T}(G)$ in search of a point $\mathcal{M}(G^*)$ where the required gluing can be done by Möbius transformations. If we can find such a point, bingo! The corresponding $\mathcal{M}_\tau$ is a hyperbolic manifold.

To simplify notation let's assume that there are just two boundary components S_1, S_2 and τ interchanges them. Consider the composed mapping which is an orientation preserving quasiconformal mapping:

$$\sigma \circ \tau : (S_1, S_2) \xrightarrow{\tau} (S_2, S_1) \xrightarrow{\sigma} (S_2', S_1').$$

It determines a homeomorphism of $\text{Teich}(S_1) \times \text{Teich}(S_2)$ onto itself. Suppose we can find a fixed point $\mathcal{M}(G^*) \in \mathfrak{T}(G)$. At this point $\sigma \circ \tau = \text{id}$ and $(S_2', S_1') \equiv (S_1, S_2)$.

Consider the meaning in $\Omega(G^*)$. Choose a component Ω_1 over S_1 and Ω_2 over S_2. In terms of lifted maps we have for $i = 1, 2$,

$$\Omega_i \xrightarrow{\tau} \Omega_j \xrightarrow{\sigma} \mathbb{S}^2 \setminus \overline{\Omega}_j.$$

Since $G_i = \text{Stab}(\Omega_i)$ is quasifuchsian, we can assume τ is a quasiconformal mapping defined on $\mathbb{S}^2$ that induces an isomorphism $\phi : G_i \to G_j$. The skinning map σ then commutes with G_j. Expressed in a different way, we have a quasiconformal map F

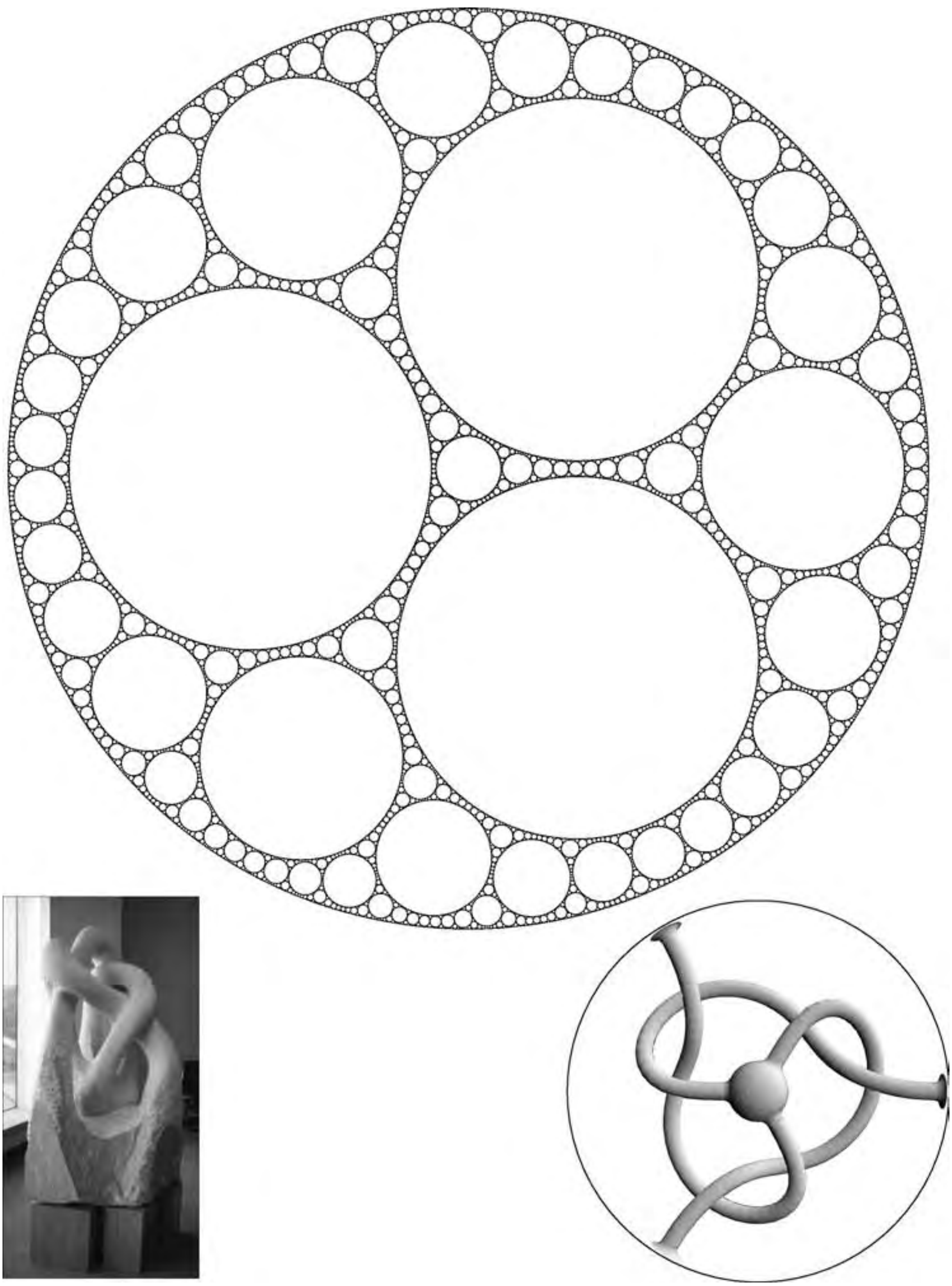


Fig. 6.1. This limit set is a Sierpiński gasket. The boundary of $\mathcal{M}(G)$ is a totally geodesic closed surface of genus two. $\mathcal{M}(G)$ itself, explained in [Thurston 1997, p. 133–138], can be embedded in $\mathbb{S}^3$, with complement a handlebody: it is what remains of an apple once a three-legged wormhole — a knotted Y shape — is eaten out (bottom right). Helaman Ferguson's marble *Knotted Wye*, standing six feet tall, was inspired by this example of Thurston. The sculpture was commissioned for the Geometry Center at the University of Minnesota, and now adorns the University's mathematics library.

defined on $\mathbb{S}^2$ with

$$F : \Omega_1 \to \mathbb{S}^2 \setminus \overline{\Omega}_2, \quad \mathbb{S}^2 \setminus \overline{\Omega}_1 \to \Omega_2.$$

Moreover F induces the isomorphism $\phi : G_1 \to G_2$. Under the hypothesis that $\mathcal{M}(G^*)$ is a fixed point, F is in fact conformal and therefore Möbius .

We claim that the augmented group $H^* = \langle G^*, F\rangle$ is discrete and $\mathcal{M}(H^*)$ is homeomorphic to M_τ, the result of identifying the boundary components S_1, S_2 via τ. This could be a job for the Klein–Maskit combination theory. Instead, we will apply Theorem 5.10.1 in Exercise 5-1:

There is a nonsingular surface $C^* \subset \mathbb{H}^3$ over Ω_1, which is invariant under G_1. Its orbit under H^* consists of mutually disjoint surfaces. Its projection $C_1 = \pi(C^*)$ into $\mathcal{M}(G^*)$ is embedded and parallel to the boundary component $S_1 = \Omega_1/G_1$. The projection $C_2 = \pi(F(C^*))$ is likewise embedded and is parallel to $S_2 = \Omega_2/G_2$. Let M denote the result of removing from $\mathcal{M}(G^*)$ the regions between C_1 and S_1, C_2 and S_2. The Möbius transformation F maps the space outside C^* onto the space inside $F(C^*)$; it effects identification of the boundary components C_1 and C_2 of M.

Assume $\mathcal{M}$ is a finite union of geometrically finite manifolds. Let $\tau : \partial\mathcal{M} \to \partial\mathcal{M}$ be an orientation reversing involution preserving the pairs of punctures on $\partial\mathcal{M}$. Form the new 3-manifold $M_\tau = \mathcal{M}/\sim\tau$ and assume that it is connected. Within $\mathcal{M}$ each puncture is paired with another by a solid pairing cylinder. Under τ the various boundary components of the pairing cylinders become identified resulting in a number of mutually disjoint boundary tori $\{T\} \subset M_\tau$. (Any solid cusp tori in $\mathcal{M}$ are left alone.) In a hyperbolic structure on $\mathcal{M}_\tau$ these must become cusp tori. To this end, we say M_τ is *atoroidal* if every map of a torus into M_τ is homotopic to a map into a component of $\{T\}$ (see Section 6.3). In particular, M_τ is atoroidal if $\mathcal{M}$ has no essential cylinders at all, that is, if $\mathcal{M}$ is acylindrical.

The key result below is proved by showing $\tau\circ\sigma$ is contracting as a homeomorphism of the relevant product Teichmüller spaces, and uniformly contracting only when M_τ is atoroidal. In the latter case it has a unique fixed point. The nontrivial proof is a job for complex analysis. For a quasifuchsian manifold $\mathcal{M}(G)$ the skinning method does not work, since σ is an isometry of the Teichmüller space. This is why a separate analysis is required. We state the result for a single manifold $\mathcal{M}(G)$:

The Skinning Lemma [Thurston 1980; McMullen 1990]. (See also[Kapovich 2001; Otal 1998].) *Assume that $M(G)$ is geometrically finite and boundary incompressible. The automorphism $\sigma\tau : \mathfrak{T}(G) \to \mathfrak{T}(G)$ has a fixed point if $\mathcal{M}(G)$ is acylindrical. The fixed point is unique.*

More precisely, the gluing problem required to construct a hyperbolic structure on M_τ has a (unique) solution if and only if M_τ is atoroidal.

The proof fails precisely when τ matches up the boundary components of an essential cylinder (not a pairing cylinder) in $\mathcal{M}(G)$ resulting in an essential torus. If $\mathcal{M}(G)$ is already acylindrical, this possibility does not occur. The presence or absence of rank two cusps is immaterial.

Actually a more general construction is needed to prove the Hyperbolization Theorem. We have to allow an incompressible subsurface of one component of $\partial\mathcal{M}(G)$ to be identified with an incompressible subsurface of another. This is carried out using a technique involving the reflection of $\mathcal{M}(G)$ over subsurfaces of $\partial\mathcal{M}(G)$. It is explained in Exercise 6-12.

6.3 The Hyperbolization Theorem

We will begin by reviewing our list of definitions and at the same time adding some new ones. For technical details see [Hempel 1976; Jaco 1980]. Suppose M^3 is a compact, orientable 3-manifold, possibly with boundary. The technical disclaimer is that we are implicitly working in the piecewise linear or equivalently in the differentiable category (see [Hempel 1976, p. 4]). The point is that we do not want to deal with "wild" embeddings. The following definitions are made with this understanding. (Here, as is customary, the terms "2-sphere" and "open or closed 3-ball" in a 3-manifold mean homeomorphic images of the round versions in $\mathbb{R}^3$. By the Sphere Theorem, the statement that M^3 is irreducible is equivalent to the condition that $\pi_2(M^3) = 0$, since the Poincaré Conjecture is now known to be true (Perelman; see [Cao and Zhu 2006]). Also, as elsewhere in this book, we are assuming that embedded surfaces S are two-sided, hence orientable: There is an embedding $\mathcal{E} : S \times [-1, 1] \hookrightarrow \mathcal{M}(G)$ with $\mathcal{E}(\partial\mathcal{M}(G) \cap S \times [-1, 1]) = \mathcal{E}(\partial S \times [-1, 1])$, if $\partial S \neq \varnothing$, and $E(x, 0) = x,\ x \in S$. See [Hempel 1976, Chapter 6; Jaco 1980, §III.12].)

Irreducible: Every embedded 2-sphere in M^3 bounds a closed 3-ball.

Boundary incompressible: If $\gamma \subset \partial M^3$ is homotopic to a point in M^3, it is already homotopic to a point in ∂M^3.

Incompressible surface: A compact, embedded surface $S \neq \mathbb{S}^2 \subset M^3$, $S \cap \partial M^3 = \partial S$ if $\partial S \neq \varnothing$, such that if $\gamma \subset S$ is homotopic to a point in M^3, it is already homotopic to a point in S. Also no component of ∂S is homotopic to a point in ∂M^3. In particular S may be a disk.

Atoroidal: Every map of a torus into M^3 that is injective on its fundamental group is homotopic to a map to a torus boundary component of M^3.

Pared manifold: $(M^3; P)$ is pared if $P \subset \partial M^3$ is the union of a finite number of mutually disjoint, incompressible annuli and tori, such that (i) every incompressible cylinder with both boundary components in P can be homotoped (relative to its boundary) into P, and (ii) every torus component of ∂M^3 is incompressible and included in P.

Acylindrical: $(M^3; P)$ has the property that $\partial M^3 \setminus P$ is incompressible and every incompressible cylinder $C \subset M^3$ bounded by simple loops $C \cap (\partial M^3 \setminus P) = \partial C$ can be homotoped (relative to ∂M^3) into ∂M^3.

Haken: M^3 is a Haken manifold if it is compact, orientable, irreducible and contains an incompressible surface $S \neq \mathbb{S}^2$ with $\partial S \subset \partial M^3$, if $\partial S \neq \varnothing$.

In an atoroidal manifold M^3, every embedded, incompressible torus is parallel to a boundary component. The following example shows that this condition alone does not suffice: R is a closed surface of genus two or the 3-punctured sphere and α is a figure-8 loop in R. In the 3-manifold $R \times \mathbb{S}^1$, the immersed torus $\alpha \times \mathbb{S}^1$ has fundamental group $\mathbb{Z} \oplus \mathbb{Z}$.

The concept of "pared" manifold arises to characterize the "parabolic loci" in those hyperbolic manifolds made compact by removing solid cusp tori and solid pairing tubes. As by paring an apple we remove blemishes on its skin, so by paring a manifold we mark the places occupied by pesky parabolics.

A compact, orientable, irreducible manifold M^3 is Haken if $\partial M^3 \neq \varnothing$. If $\partial M^3 = \varnothing$, then M^3 is Haken if and only if (i) the cardinality of the first homology group is infinite and/or (ii) $\pi_1(M^3)$ is a free group with amalgamation over a closed surface group or is an HNN-extension [Waldhausen 1968, Lemma 1.1.6]. Condition (ii) implies that there is an embedded incompressible surface not parallel to a boundary component; the converse is true by van Kampen's theorem. Haken's original term for the class of Haken manifolds, "sufficiently large", is suggestive of what is required of the fundamental group, although no precise characterization in terms of this group is known. A criterion in distinguishing between Haken and non-Haken closed manifolds that has been useful is the following curious theorem of Hyman Bass:

Theorem 6.3.1 [Bass 1980; Maclachlan and Reid 2003, Corollary 5.2.3]. *Suppose* $\mathcal{M}(G)$ *has finite volume. Either* $\mathcal{M}(G)$ *contains a closed incompressible surface not homotopic to a cusp torus, or* $G \subset \mathrm{SL}(2, \mathbb{C})$ *is conjugate to a subgroup of* $\mathrm{SL}(2, \mathbb{A})$, *where* $\mathbb{A}$ *is the ring of algebraic integers in the algebraic closure* $\overline{\mathbb{Q}}$.

The importance of the class of Haken manifolds arises from the fact that if M^3 is Haken, it has a "hierarchy"

$$M^3 = M_1 \supset M_2 \supset \cdots \supset M_n = \mathbb{B}^3 \big(\bigcup \mathbb{B}^3\big).$$

Here $S_k \subset M_k$ is an orientable, incompressible, nonseparating (for $k > 1$) surface, and

$$M_{k+1} = M_k \setminus S_k.$$

The surface S_k is not boundary parallel and it is properly embedded. Possibly S_k has a boundary in which case $S_k \cap \partial M_k = \partial S_k$ a union of incompressible simple loops. If M^3 is closed, then the first surface may disconnect M^3 resulting at the end of the decomposition in two balls. For details see [Hempel 1976; Jaco 1980].

M^3 can be systematically so decomposed by a sequence of incompressible surfaces; equally it can be systematically composed from one or two 3-balls by successively forming M_k from M_{k+1} by identifying a pair of disjoint incompressible surfaces or subsurfaces on the boundary of M_{k+1} and gluing them together to form M_k. If M_{k+1} is hyperbolic, the resulting M_k can be made hyperbolic as well by an application of the Skinning Lemma (page 321) — if full boundary components are to be joined; otherwise an elaboration is needed (see Exercise 6-12).

Hyperbolization of 3-Manifolds. *Assume that M^3 is compact, orientable, irreducible, atoroidal, and pared* $(M^3; P)$. (Possibly $P = \varnothing$; see Exercise 6-1.)

(i) *Suppose $\partial M^3 \neq \varnothing$. Then $M^3 \setminus P$ is homeomorphic to $\mathcal{M}(G)$ for some geometrically finite group G. $\mathcal{M}(G)$ is compact if and only if $P = \varnothing$; $\mathcal{M}(G)$ has finite volume if and only if $\partial M^3 = P$ consists of incompressible tori.*
(ii) *If M^3 is closed and Haken, it is homeomorphic to a $\mathcal{M}(G)$.*

For an overview see [Thurston 1980; 1982b; Morgan 1984; Bonahon 2002; Scott 1983]. For proofs see [Otal 1996; 1998; Kapovich 2001].

The classification of manifolds that can be hyperbolized is complete because the following conjecture is established as part of the Geometrization Conjecture by Perelman (see Section 6.4).

Hyperbolization Conjecture. *Assume that M^3 is closed, orientable, irreducible with infinite fundamental group $\pi_1(M^3)$. Then M^3 is homeomorphic to a kleinian manifold $\mathcal{M}(G)$ if and only if one of the following conditions hold*:

(i) *$\pi_1(M^3)$ contains no noncyclic abelian subgroup,*
(ii) *$\pi_1(M^3)$ contains no cyclic normal subgroup.*

In view of the Hyperbolization Theorem we can stick to closed manifolds. We know the conditions listed are necessary. A bridge between the two statements is Scott's Strong Torus Theorem [1980] which says, for an orientable, irreducible, compact M^3 for which $\pi_1(M^3)$ has a rank two abelian subgroup, that either M^3 contains an embedded incompressible torus or $\pi_1(M^3)$ contains a cyclic normal subgroup K. The Seifert conjecture, confirmed in [Gabai 1992, §8.6] and [Casson and Jungreis 1994], says that if $\pi_1(M^3)$ is infinite, the latter case occurs if and only if M^3 is a Seifert fibered space and then $\pi_1(M^3)/K$ is either the fundamental group of a 2-orbifold, or is fuchsian (with elliptics).

Agol's proof of tameness also establishes that noncompact, complete, orientable riemannian 3-manifolds with pinched negative sectional curvatures $-\infty < -L < \kappa < -l < 0$ and finitely generated fundamental groups have tame ends (personal communication). Therefore they too carry hyperbolic metrics, partially settling a long standing problem.

An alternate approach to hyperbolization is through the Virtual Haken Conjecture:

If M^3 is an irreducible 3-manifold with infinite $\pi_1(M^3)$, then M^3 has a finite cover which is Haken.

In [Dunfield and Thurston 2003] this was tested on the Hodgson–Weeks census of the 10,986 smallest volume closed hyperbolic manifolds, most but not all of which have finite homology. The fundamental groups all are 2- or 3-generator. In every case the conjecture was confirmed by showing each had a cover with infinite homology. The authors also show that every nontrivial Dehn surgery on the figure-8 knot complement results in a virtual Haken manifold. Whether the Haken cover is hyperbolic depends on whether it is atoroidal. The affirmation of the conjecture would imply the solution of the geometrization conjecture for closed irreducible manifolds with infinite fundamental groups since Haken manifolds are either Seifert fibered or hyperbolic.

6.3.1 Knots and links

We now cite Thurston's famous theorems for knot and link complements, clearly displaying the power of his theory.

Hyperbolic Structures on Knot Complements. *If $K \subset \mathbb{S}^3$ is a knot, $\mathbb{S}^3 \setminus K$ has a hyperbolic structure if and only if K is not a torus knot or a satellite knot.*

A torus knot is a nontrivial knot that lies on the boundary of a torus neighborhood of an unknot in $\mathbb{S}^3$. A satellite (also called a companion) knot K is one that is embedded in a small solid torus neighborhood of some knot K_0, not the unknot, and K is not isotopic to K_0 nor is contained in a ball inside the solid torus. The torus about K_0 is incompressible in $\mathbb{S}^3 \setminus K$, which makes it impossible for K to have a hyperbolic complement.

By a celebrated theorem of Gordon and Luecke [1989], knots are determined by their complements: If K_1, K_2 are two knots, there is a homeomorphism (orientation preserving or reversing) of $\mathbb{S}^3$ taking K_1 to K_2 if and only if $\mathbb{S}^3 \setminus K_1$ is homeomorphic to $\mathbb{S}^3 \setminus K_2$. (This is not true for link complements). Therefore the volume of the complement of a hyperbolic knot is an invariant of the knot.

One might think it very rare that a closed manifold with infinite fundamental group has zero first homology. Yet given a hyperbolic knot K, there are infinitely many Dehn surgeries on K which result in a closed manifold with exactly this property. To construct examples, start with tubular neighborhood $\mathcal{T}$ of K. On $T = \partial\mathcal{T}$ there is a uniquely defined (up to free homotopy) pair of simple closed curves which cross each other once: The *meridian* μ bounds a disk within $\mathcal{T}$ and generates the homology of $\mathbb{S}^3 \setminus \mathcal{T}$. The *longitude* λ is parallel to K and is homologous to zero in $\mathbb{S}^3 \setminus \mathcal{T}$. This is because there exists a Seifert surface for K — an orientable surface $S \subset \mathbb{S}^3 \setminus K$ with $\partial S = K$ [Lickorish 1997].

Do $(n, 1)$-Dehn surgery: replace the inside of the solid torus $\mathcal{T}$ by another solid torus, so that the simple loop $\lambda^n \mu \subset T$ becomes the meridian. This yields a closed manifold M. In M, both λ and λ^n are homologous to 0; hence so is μ. Consequently the first homology group of M is zero. Thurston shows that for all except a finite number of integers n, the resulting manifold has a hyperbolic structure.

In general, there is no known method to determine whether a particular 1-cusped manifold is a knot complement in the 3-sphere or in some other manifold.

In [Koundouros 2004] the following interesting conjecture is proposed and explored: If the injectivity radius of the closed manifold $\mathcal{M}(G)$ is sufficiently large, then $\mathcal{M}(G)$ cannot be obtained by Dehn surgery on a knot in $\mathbb{S}^3$. The *injectivity radius* of a closed manifold is the largest number r such that every point in the manifold is the center of an embedded ball of radius r.

A link $L \subset \mathbb{S}^3$ is called *indecomposable* if it cannot be separated into two parts which can be isotoped into disjoint balls. A satellite (or companion) link L is satellite to another link L_0 if one or more components of L is satellite to a component of L_0. Thurston proved the following two theorems. The first is recorded in [Epstein and Gunn 1991, p. 41]; the second appears in [Thurston 1982b; 1979, p. 5.38].

Hyperbolic Structures on Link complements. *Suppose $L \subset \mathbb{S}^3$ is an indecomposable link of $m \geq 2$ components. Suppose no component is a torus knot and L is not a satellite link. Then $\mathbb{S}^3 \setminus L$ has a hyperbolic structure.*

The hyperbolic structure for the Borromean rings complement is visualized in [Gunn and Maxwell 1991].

Dehn Surgeries on Hyperbolic Link Complements. *Suppose $L \subset M^3$ is a link in the 3-manifold M^3, in particular in $\mathbb{S}^3$, such that $M^3 \setminus L \cong \mathcal{M}(G)$ has a hyperbolic structure. For each cusp of $\mathcal{M}(G)$ there are a finite number of Dehn surgeries that must be excluded. The manifolds resulting from all Dehn surgeries on $\mathcal{M}(G)$, except for those excluded, have a hyperbolic structure.*

We have earlier stated this for the case of knots; see page 214. It is known that every closed 3-manifold is obtained by Dehn surgery along some link $L \subset \mathbb{S}^3$ whose complement is hyperbolic. Most of these Dehn surgeries also give rise to non-Haken manifolds [Thurston 1982b; 1979]. See also Exercise 6-3.

It is also true that every closed orientable 3-manifold is a cover of $\mathbb{S}^3$ branched over the Borromean rings [Hilden et al. 1985]. For the construction of k-fold unbranched covers of $\mathbb{S}^3 \setminus L$, see [Rolfsen 1976, §10.F].

The program SnapPea, by Jeff Weeks [2005], allows the computation of the hyperbolic structure of knots known to have hyperbolic structure and, for practical reasons, are not too complicated (see also Exercise 4-18). It turns out that, contrary to prior assumption, computation is easier in the conformal model than in the hyperboloid model [Floyd et al. 2002]. The proof that the algorithm underlying SnapPea in principle results in the correct structure is contained in an enhancement to SnapPea called Snap Goodman n.d.; Coulson et al. 2000, which can compute arithmetic invariants, such as volumes, to very high precision. Using Snap one can confirm the hyperbolic structure as discovered by SnapPea by finding the exact solutions of the equations satisfied by the tetrahedral parameters needed to construct the cusped manifold from ideal tetrahedra (if the degree is not too high). Thus it is possible, in principle, to decide whether or not a given knot, presented say by over and under crossings, is the unknot. Likewise the question of whether two hyperbolic knots are the same or not can be answered in principle when SnapPea can find a hyperbolic structure for each. See also Exercise 1-23 on computing volumes.

Riley [1975] (see also [Wielenberg 1978]) discovered a variety of kleinian groups among the Bianchi groups $\Gamma_d = \mathrm{PSL}(2, \mathcal{O}_d)$ and their finite index subgroups. Here $\mathcal{O}_d$ denotes the ring of integers in the quadratic imaginary number field $\mathbb{Q}(\sqrt{-d})$,

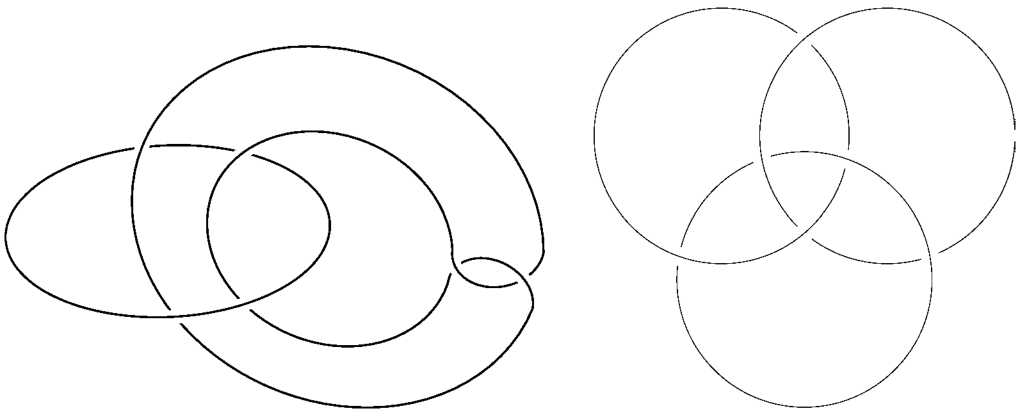


Fig. 6.2. The Whitehead link and the Borromean rings.

$\mathcal{O}_d$ denotes the ring of integers in the quadratic imaginary number field $\mathbb{Q}(\sqrt{-d})$, where d is a positive integer. Many of these model familiar knots and links. The ring $\mathcal{O}_d$ has $\mathbb{Z}$-basis $\{1, \omega\}$ where $\omega = \sqrt{-d}$, unless $d+1$ is divisible by 4 in which case $\omega = \frac{1}{2}(1+\sqrt{-d})$. So the normalized matrices in the group have entries of this form. All of the Bianchi groups, together with their finite index subgroups, give rise to manifolds of finite volume. In many cases, an exact formula can be given for the volume; see [Milnor 1994, p. 257].

In particular the Picard group of Exercise 2-11 is $\mathrm{PSL}(2, \mathcal{O}_1)$. The Borromean rings come from the one torsion-free normal subgroup of index 24 in $\mathrm{PSL}(2, \mathcal{O}_1)$ [Wielenberg 1978; Brunner et al. 1984]. The figure-eight knot complement of see Exercise 3-5 h has index 12 in $\mathrm{PSL}(2, \mathcal{O}_3)$. The Weeks manifold, the conjectured lowest volume (orientable) hyperbolic manifold, is obtained from $\mathbb{Q}(x)$ with $x^3 - x + 1 = 0$. The field $\mathbb{Q}(x)$ where $x^4 - x^2 + 3x - 2 = 0$ spawns a closed manifold all of whose geodesics are simple—in fact an infinite family of such manifolds has been found! The definitive reference on this subject of *arithmetic kleinian groups* is [Maclachlan and Reid 2003]. There the reader will also find an exhaustive list of known examples. See also [Riley 1975] and [Thurston 1979, Chapter 7].

There is another approach to knot invariants, namely via the many known knot polynomials. It is not known how to find these polynomials directly from the hyperbolic structure.

In general, there is no known method to determine whether a particular 1-cusped manifold is a knot complement in the 3-sphere or in some other manifold.

6.4 Geometrization

There is a general conjecture of grand sweep proclaimed in 1977 by William Thurston, called the *geometrization conjecture*. It has been the focus of 3-manifold topology ever since. When he proposed the conjecture, Thurston announced the solution of a substantial part of it. If completely proved, it would amount to a complete classification of compact 3-manifolds, not just a topological classification but in fact a geometric classification. For many explicit examples of 3-dimensional manifolds and orbifolds see [Montesinos 1987].

It had been known that there is a process of canonically cutting down a compact, orientable 3-manifold into compact pieces by embedded spheres and two-sided incompressible tori. First cut along an embedded sphere that does not bound a ball and cap off the resulting two pieces. After a finite number of steps the process will stop. The summands in the decomposition are unique up to homeomorphisms; see [Hempel 1976; Jaco 1980]. Then for each irreducible piece there is the Johannson–Jaco–Shalen decomposition by a finite set of mutually disjoint, embedded, incompressible tori with the following properties: none of the tori is parallel to a boundary component, and each component resulting from cutting the manifold along the tori is either a *Seifert fiber space* or it contains no incompressible torus. A minimal set of tori is uniquely determined up to isotopy; see [Jaco 1980].

(An example of a Seifert fiber space, or *Seifert manifold*, is given by $M^3 = R \times \mathbb{S}^1$, with R a compact surface; the boundary components of R, if any, become incompressible boundary tori for $\mathcal{M}^3$. Other examples are obtained by replacing a finite number of circles in M^3 with "singular fibers"; a singular fiber has a neighborhood homeomorphic to the quotient $\mathbb{D}\times\mathbb{S}^1$ of $\mathbb{D}\times\mathbb{R}$ under the action $(z, t)\mapsto(\omega z, t+1/q)$, where ω is a primitive q-th root of unity and $\mathbb{D}$ is the open unit disk centered at $z=0$. Each nonsingular fiber wraps q-times around the singular one. An orientable M^3 is called Seifert fibered if it is a union of pairwise disjoint simple loops, each with a closed neighborhood, a union of fibers, which is fiber-homeomorphic to a fibered solid torus $\mathbb{D}\times\mathbb{S}^1$ as described above; see [Jaco 1980; Scott 1983]. The only Seifert fibered manifolds that appear inside hyperbolic manifolds are interiors of solid tori or solid cusp tori.)

The geometrization conjecture is the statement that the interior of each resulting submanifold has a uniquely determined geometric structure. Here "geometric structure" means the following: Assume that X is a simply connected, complete riemannian manifold which is homogeneous — there is an isometry taking any point to any other. Assuming X is a 3-manifold, it is diffeomorphic to $\mathbb{S}^3$ or $\mathbb{R}^3$, unless it is modeled on $\mathbb{S}^2\times\mathbb{R}$. One can say that a complete riemannian manifold M is modeled on (X, g) if $M = X/G$, where G is a fixed point free group of isometries.

Thurston conjectured that the interior of each (compact) piece is modeled on one of eight kinds of geometries (see [Thurston 1997, Thm. 3.8.4]). The most familiar are the constant sectional-curvature geometries: spherical, euclidean and hyperbolic. There are five other geometries possible as well: $\mathbb{S}^2\times\mathbb{R}$, $\mathbb{H}^2\times\mathbb{R}$ (both of which have product metrics), $\widetilde{\mathrm{SL}(2,\mathbb{R})}$ (the universal covering group of $\mathrm{SL}(2,\mathbb{R})$), Nil and Sol [Thurston 1982b; Scott 1983; Boileau et al. 2003]. The formerly unresolved cases in the geometrization conjecture fall into two types according to the nature of the fundamental group: (i) $\pi_1(M^3)$ is finite; in this case the conjecture is that $M^3\cong\mathbb{S}^3/\Gamma$

where Γ is a finite orthogonal group acting in $\mathbb{R}^4$ (the Poincaré Conjecture is the case $\Gamma = \mathrm{id}$); (ii) $\pi_1(M^3)$ does not contain a cyclic normal subgroup; this is the Hyperbolization Conjecture.

The confirmation of the Seifert Conjecture by Gabai [1992] and independently by Casson and Jungreis [1994] settled the issue for manifolds for which $\pi_1(M^3)$ does contain a cyclic normal subgroup. Namely if M^3 is compact, orientable, irreducible with infinite $\pi_1(M^3)$, then M^3 is a Seifert fibered space if and only if $\pi_1(M^3)$ contains a cyclic normal subgroup. A Seifert fibered space cannot have hyperbolic interior, unless it is the elementary case of a solid torus or thickened torus (solid torus minus its core).

As our book nears completion, the mathematical world seems to have confirmed Grigori Perelman's announcement of having completed the proof of the Geometrization Conjecture for closed manifolds, including the Poincaré Conjecture. The first full proof independently confirming Perelman's announcement is by H.-D. Cao and X.-P. Zhu [2006]. The strategy of the ultimately successful proof was set out by Richard Hamilton some years ago: Allow the Ricci flow to act on a given 3-manifold and deal with the singularities as they arise, pinching off pieces of the manifold. For expositions see [Chow and Knopf 2004; Morgan 2005; Milnor 2003].

For interesting comparisons of hyperbolic volumes and volumes of riemannian 3-manifolds, assuming Perelman's work, see [Agol et al. 2005].

In the context of the geometrization conjecture, one can proclaim that the "vast majority" of compact 3-manifolds are hyperbolic, although there is not yet a formal theorem making matters precise. Compare the manifold case with the case of word-hyperbolic groups (Exercise 2-17.

6.5 The Orbifold Theorem

As we have discussed in Section 2.5, if the kleinian group G has elliptic elements, its quotient $\mathcal{M}(G)$ is called an (orientable) *orbifold*. Assume that $\mathcal{M}(G)$ is geometrically finite. We will collect the properties of its singular set $\sigma(G)$, which is a graph, as follows — Corollary 2.5.2:

(i) Each edge e is labeled by a positive integer $k \geq 2$ which is the order of a primitive elliptic element that pointwise fixes a lift $e^* \subset \mathbb{H}^3$.

(ii) Each vertex of $\sigma(G)$ is the endpoint of three edges. These edges must have orders $(2, 3, 3)$, $(2, 3, 4)$, $(2, 3, 5)$ or $(2, 2, n)$, $n \geq 2$.

(iii) A component of $\sigma(G)$ is either a simple loop, a full geodesic with each endpoint in $\partial\mathcal{M}(G)$ or at a cusp, or it contains vertices. An edge from a vertex ends either at another vertex, a point on $\partial\mathcal{M}(G)$, or at a cusp.

(iv) If an edge ends at a rank one cusp, there are exactly two such edges and each has the label 2.

(v) If an edge ends at a rank two cusp, there are three or four such edges with the labels $(3, 3, 3)$, $(2, 3, 6)$, $(2, 4, 4)$, or $(2, 2, 2, 2)$.

(vi) There are no euclidean or spherical suborbifolds except as arise automatically from a cusp or a fixed point of a finite subgroup.

A spherical submanifold means that there is a smooth topological sphere S that is transverse to $\sigma(G)$, cuts each edge at most once, and doesn't intersect the vertices. Set $S \cap \sigma(G) = \{p_i\}$. Each point p_i has a label r_i coming from the edge it lies on. In view of Equation 3.13, we must have $\sum(1 - 1/r_i) > 2$.

For a definition of orientable orbifold see Exercise 4-7; for the general definition see [Cooper et al. 2000; Kapovich 2001, Chapter 6; Thurston 1979, Chapter 13]. Thurston proclaimed a complete geometrization theory for compact, irreducible, atoroidal orbifolds. Proofs covering many cases have been developed by Boileau, Leeb, and Porti [Boileau et al. 2005], and independently by Cooper, Hodgson, and Kerckhoff [Cooper et al. 2000]. A proof when the singular set is a union of circles is in [Boileau and Porti 2001]. The proofs require analysis of cone manifolds with cone angles $\leq \pi$ when trivalent vertices are present.

Instead of plunging into the general theory we will be more than content to present a beautiful application to hyperbolic manifolds:

Designated Hyperbolic Orbifolds. *Suppose $\mathcal{M}(G)$ is a geometrically finite hyperbolic manifold, and $\{\gamma_i\}$ is a finite set of mutually disjoint, simple closed geodesics with assigned integers $\{2 \leq r_i \leq \infty\}$. There exists a hyperbolic orbifold $\mathcal{M}(H)$ such that the labeled graph $\sigma(H)$ is homeomorphic to $\bigcup \gamma_i$ with the given labeling and $\mathcal{M}(H) \setminus \sigma(H)$ is homeomorphic to $\mathcal{M}(G) \setminus \bigcup \gamma_i$.*

Proof. This is a direct consequence of [Boileau et al. 2005, Theorem 2.3]. The presence of the graph determines orbifold structure $\mathcal{O}$ with "base space" $\mathcal{M}(G)$. It is topologically atoroidal, not a Seifert fibered orbifold, nor is there a spherical or euclidean suborbifold. For more details see Exercise 6-3. □

There are many other possibilities for a hyperbolic orbifold besides the ones stated above, but the analysis is more complicated. One problem is that a singular set cannot consist of a "knotted" geodesic. Another is that the underlying manifold to a hyperbolic orbifold may or may not have a hyperbolic structure. The example of $\mathbb{H}^2$ covering $\mathbb{S}^2$ with branch points satisfying Equation 3.11 or 3.12 shows why. It requires the extensive analysis of [Boileau et al. 2005; Cooper et al. 2000] to explore all the possibilities.

Suppose $\mathcal{M}(G)$ is a closed manifold. We can ask, of all the hyperbolic orbifolds whose underlying space is $\mathcal{M}(G)$, which has the least volume? Peter Storm gave the following answer, which is consistent with the Dehn Surgery Theorem of page 214 (see also Theorem 6.6.5):

Theorem 6.5.1 [Storm 2002b]. *The volume of the closed manifold $\mathcal{M}(G)$ is strictly less than the volume of any hyperbolic orbifold $\mathcal{O}$ (or, more generally, any cone manifold $\mathcal{O}$ with cone angles $\leq 2\pi$) with the property that the underlying space of $\mathcal{O}$ is homeomorphic to $\mathcal{M}(G)$.*

ifold $\mathcal{O}$ with cone angles $\leq 2\pi$) with the property that the underlying space of $\mathcal{O}$ is homeomorphic to $\mathcal{M}(G)$.

6.6 Exercises and Explorations

6-1. The interior of $M^3 = \mathbb{T}^2 \times [0, 1]$ resulting from thickening the torus $\mathbb{T}^2$ with $P = \mathbb{T}^2 \times \{0\}$ would fall under the aegis of the Hyperbolization Theorem (page 324), except that it is not pared since there are essential cylinders with one boundary component on P. Its interior is homeomorphic to $\mathbb{R}^3/\Gamma_0$, where Γ_0 is the group generated by $\langle (x, y, t) \mapsto (x + 1, y, t),\ (x, y, t) \mapsto (x, y + 1, t) \rangle$. However its interior does not have a complete hyperbolic structure of *finite volume*; this makes it the sole exception to the rule that (finitely generated) hyperbolic manifolds all of whose ideal boundary components are tori have finite volume.

Consider instead the orientation and volume preserving group Γ of euclidean isometries of $\mathbb{R}^3$ generated by $(x, y, t) \mapsto (x + 1, y, t)$, $(x, y, t) \mapsto (-x, y + 1, -t)$. The group Γ preserves $\mathbb{C}$ and $\mathbb{C}/\Gamma$ is a Klein bottle; the torus $\mathbb{C}/\Gamma_0$ is its two sheeted orientable cover. The corresponding manifold $\mathbb{R}^3/\Gamma$ is homeomorphic to the interior of the compact manifold with torus boundaries $N^3 = \mathbb{T}^2 \times [0, 1]/(\mathbb{Z}/2)$. N^3 itself is called the *twisted I-bundle* over the Klein bottle (see [Hempel 1976, Theorem 10.3]). Confirm that N^3 does not have a pared structure either. Does its interior have a complete hyperbolic structure? (Answer: No!)

6-2. *Doubling a manifold.* The following construction leads to another proof of Theorem 6.2.1. Suppose $M = \mathcal{M}(X)$ is geometrically finite, acylindrical, without cusps, and with $b \geq 1$ boundary components.

Here is a process for constructing a *closed* manifold $\widehat{M}$ that is 2^b-sheeted over M and is orientable, atoroidal, and Haken. In the case $b = 1$ it is just the double of M.

Order the boundary components of $M : (S_1, S_2, \dots)$. Take 2^b copies of M. On each copy, assign to each of its boundary components one of the two symbols $+$ and $-$. We do this so that the ordered boundary of each copy is associated with a symbol sequence and no two copies are assigned the same ordered symbol sequence. Index the copies as M_i.

Glue M_i to M_j along the surface S_k if and only if the symbol sequence for M_i differs from that of M_j only in the k-th position. Think of this operation as M_j being the reflection of M_i across S_k. If we start with a manifold with totally geodesic boundaries, we can carry out this operation in the universal cover $\mathbb{H}^3$.

There is a finite group G of order 2^b of automorphisms acting on $\widehat{M}$. We can see this by considering the connected b-valent graph $\mathcal{G}$ constructed as follows. The vertices correspond to the copies $\{M_i\}$. Two vertices are joined by an edge if one manifold is glued to the other. Denote by J_k the involution of $\mathcal{G}$ determined by replacing the sign attached to the k-th boundary component of each M_i by the opposite sign. Each J_k acts on $\widehat{M}$ and is an orientation reversing involution: "reflection" in each S_k. Denote the group generated by the $\{J_k\}$ by G. No element of G fixes a vertex of $\mathcal{G}$ (other than the identity), and for each index i, there is an element w_i taking the vertex corresponding to M_1 to that corresponding to M_i. The expression for w_i as a word in the generators is not uniquely determined. Rather what is determined is the parity of its length, even or odd, orientation preserving or reversing.

The original M appears as the quotient space $\widehat{M}/G$. Hyperbolize $\widehat{M}$ to become the closed manifold $\mathcal{M}(Y)$. For example by Theorem 3.13.1, there is a homeomorphism $f : \widehat{M} \to \mathcal{M}(Y)$. Moreover each element of fGf^{-1} is isotopic to an isometry by Mostow's theorem and G is isomorphic to a group H of isometries. The boundary components of $\mathcal{M}(Y)/H \cong M$ are totally geodesic.

6-3. *Drilling out simple geodesics.* Start with a geometrically finite hyperbolic manifold $\mathcal{M}(G)$ with boundary. The result of removing a simple loop which is *not* primitive cannot have a hyperbolic structure (Exercise 4-20) but simple geodesics are automatically primitive. The result of removing a finite number of mutually disjoint simple geodesics is also atoroidal and Haken and consequently has a hyperbolic structure. On the other hand, if $\mathcal{M}(G)$ has finite volume with cusps, a closed hyperbolic manifold can be obtained by Dehn surgeries. The original manifold can in turn be recovered by removing the resulting new geodesics and introducing the original hyperbolic structure on the complement. The general theorem covering this matter is as follows.

Theorem 6.6.1 (Kerckhoff, Kojima, Sakai). *If $\mathcal{M}(G)$ is geometrically finite and $L \subset \mathcal{M}(G)$ is a disjoint union of a finite number of simple closed geodesics, then $\mathcal{M}(G) \setminus L$ has a complete hyperbolic structure.*

As a consequence of the Dehn Surgery Theorem (p. 214), the volume of $\mathcal{M}(G)$, if it does have finite volume, is less than that of the hyperbolic structure on $\mathcal{M}(G) \setminus L$.

We will assume that $\mathcal{M}(G)$ is closed and indicate the proof as in [Kojima 1988, Prop. 4]. First, $\mathcal{M}(G) \setminus L$ is irreducible — every embedded 2-sphere bounds a ball. (Hint: lift to $\mathbb{H}^2$.) Nor does it contain an essential embedded torus T which cannot be homotoped into L. (Hint: Suppose there is an essential embedded torus T. With respect to $\mathcal{M}(G)$, T is compressible so there is a compressing disk D. Choose a lift $\widehat{T}$ of T to $\mathbb{H}^3$. Then $\widehat{T}$ is either a torus or an open cylinder. In the former case, the region bounded by $\widehat{T}$ could not contain a lift of any of the geodesic components of γ. Therefore $\widehat{T}$ would be contractible in $\mathbb{H}^3 \setminus L^*$, where L^* is the totality of lifts of L. As such T would be contractible in $\mathcal{M}(G) \setminus L$, a contradiction. If the second case occurred, $\widehat{T}$ would be invariant under a cover transformation $g \in G$, and hence its closure would have two ideal points on $\partial\mathbb{H}^3$, the fixed points of g. $\widehat{T}$ incloses a cigar-shaped region which would contain exactly one component of L^*, the lift of some component L_0 of L. In this case T is the boundary of a tubular neighborhood of L_0.)

It remains to show that $M' = \mathcal{M}(G) \setminus L$ cannot be a Seifert fiber space (SFS) — see Section 6.4 or [Scott 1983] for the formal definition. If to the contrary it were a SFS, M' would have the following structure (mandated by the Torus Theorem — see [Jaco 1980]). There would be a fuchsian group H with possible elliptic elements such that $S = \mathbb{H}^2/H$ is a closed surface and $\pi_1(M')/K \cong H$, where K is an infinite cyclic normal subgroup. The singular fibers come from the elliptic fixed points. Since $\mathcal{M}(G)$ is formed from M' by adding back L, there would be at most one singular fiber and that would have to be a meridian of T, because $\mathcal{M}(G)$ is not a SFS. Therefore we would have $G \cong H$. Now G is torsion free so H must be as well and hence is a surface group. This is impossible as the third homology of S is zero while that of $\mathcal{M}(G)$ is not.

The argument presented was kindly provided by Peter Scott.

Can you hyperbolize $\mathcal{M}(G) \setminus \{\gamma\}$ where γ is a simple loop in the interior which does not bound an embedded disk in its complement?

Under some circumstances it may be advantageous not to globally change the hyperbolic structure because there is no control of how the geometry changes. Instead one can replace the hyperbolic metric near γ by a complete PNC metric. What is needed is the following result, formulated by Souto (see also [Hodgson and Kerckhoff 1998]):

Theorem 6.6.2. *Given $\varepsilon > 0$ there exists $C > 0$ with the following property. If $\gamma \subset \mathcal{M}(G)$ is a simple closed geodesic such that its R-tube V is embedded, then the hyperbolic metric on $\mathcal{M}(G) \setminus V$ can be extended to a complete PNC metric on $\mathcal{M}(G) \setminus \gamma$ which in V has curvature κ satisfying $-1-\varepsilon < \kappa < -1+\varepsilon$.*

It is shown in [Agol 2002] that if a closed geodesic γ is removed from a hyperbolic 3-manifold $\mathcal{M}(G)$ of finite volume, and if γ is contained in an embedded tube of radius r, then the volume of the hyperbolic structure M_γ on $\mathcal{M}(G) \setminus \gamma$ satisfies

$$\mathrm{Vol}(M_\gamma) \leq (\coth r)^{5/2}(\coth 2r)^{1/2}\mathrm{Vol}(\mathcal{M}(G)).$$

Compare Theorem 6.6.1 with the following striking result Robert Myers [1982]:

Theorem 6.6.3. *Let M be a compact, orientable 3-manifold for which ∂M, if nonempty, does not contain a 2-sphere. Then M contains a simple closed curve K whose open tubular neighborhood N_K has the following property. The complement $M \setminus N_K$ is irreducible, boundary incompressible, and atoroidal.*

In particular $\mathrm{Int}(M) \setminus K$ has a complete hyperbolic structure. For example if $M \cong S \times [0, 1]$ for a closed surface S, there is a knot K so that $M \setminus K$ is acylindrical. For this to happen the projection of K to S must fill up S in the sense that each complementary component is simply connected.

Meyers goes on to prove that M is completely determined as a 3-manifold by the countably infinite set of subgroups $\{\pi_1(M \setminus J)\}$ as J runs over all simple loops in M.

Instead of drilling out a geodesic we turn to the following, related situation. Suppose instead of a manifold we have a hyperbolic cone manifold M (Exercise 4-7) such that the singular locus c is a simple closed curve with cone angle α. Still M may have a geometrically finite structure and in particular a conformal boundary as befits geometrically finite hyperbolic manifolds. The following theorem of Ken

Bromberg [2004], proved using work of Craig Hodgson and Steve Kerckhoff has been of fundamental importance in applications.

Theorem 6.6.4. *Suppose M_α is a geometrically finite hyperbolic cone manifold without rank one cusps with cone angle $0 \le \alpha \le 4\pi$. Suppose the singular locus c has a tubular neighborhood of radius at least* $\operatorname{arcsinh}(\sqrt{2})$. *There exists $\varepsilon = \varepsilon(\alpha)$ such that if the length of c satisfies $L_\alpha(c) < \varepsilon$ then there exists a family of geometrically finite cone manifolds $\{M_t\}$ with cone angle t, $0 \le t \le \alpha$, and conformal boundary fixed.*

Thus we can continuously decrease the cone angle, for example from the initial 4π to 2π, at which point $M_{2\pi}$ is a complete hyperbolic manifold, or to 0 when c becomes a rank two cusp, which is the case of Theorem 6.6.1. Cone angles of 4π arise naturally from the construction of Exercise 6-7.

The change in geometry of the manifolds as t varies is controlled by the following theorem. We fix neighborhoods U_t of the cone axis in M_t by taking the appropriate component of the ε-thin part for a sufficiently small ε. As $t \to 0$, U_t approaches a horoball.

Drilling Theorem 6.6.5 [Brock and Bromberg 2003]. *Suppose M_α satisfies the hypothesis of Theorem 6.6.1. Given $L > 1$, there exists $\varepsilon = \varepsilon(\alpha; L)$ with the following property. For $L_\alpha(c) < \varepsilon$ and $0 \le t \le \alpha$, there is an L-bilipschitz map*

$$M_\alpha \setminus U_\alpha \to M_t \setminus U_t$$

that sends $\partial U_\alpha \to \partial U_t$.

This result is quite strong: it says in quantitative terms that the manifolds $\{M_t\}$ remain a bounded distance apart as $t \to 0$; when a simple geodesic is drilled out, away from the geodesic the hyperbolic structure does not change much. The closer L is to 1, the closer c must be to a rank two cusp. The two theorems have been applied in the study of the density of geometrically finite groups, the density of cusps on boundaries of deformation spaces, the Ending Lamination Conjecture, and tameness of manifolds. The theorems have full analogues in case of a finite number of mutually disjoint simple cone axes (Bromberg, personal communication; it seems likely that the restriction on rank one cusps is not necessary — although one can make it a rank two cusp by adjoining another parabolic).

One application is to the proof of Theorem 4.6.3. In an approximating sequence one drills out the short geodesics which are destined to become the rank one parabolics in the algebraic limit. This does not change the hyperbolic structure away from these geodesics very much. Consider the covering manifolds determined by the marked fundamental groups of the sequence of drilled out cores. These covers converge algebraically to the expected algebraic limit.

Another application is in the uniqueness proof of Bonahon and Otal's Theorem 3.11.3. If we have a convex core whose bending lamination is finite, we can construct the double of the convex core across its boundary components. If a particular bending line ℓ has bending angle α so that the internal bending angle is $\pi - \alpha$, then the result of doubling gives a cone manifold with cone axis ℓ and cone angle $2(\pi - \theta)$ and similarly the other bending lines become cone axes as well. The manifold can be continuously deformed until all the cone angles become zero. Then Mostow rigidity can be applied.

6-4. *Knottedness.* What should it mean that a simple geodesic γ in $\mathcal{M}(G)$ be *unknotted*? In $\mathbb{S}^3$, a simple loop is unknotted if it is isotopic to a point in its complement. We could also define a simple loop to be unknotted if it isotopic to a simple loop in a sphere.

One definition might be as follows. Consider the countable set in $\mathbb{H}^3$ of all lifts of γ. Call γ unknotted if any geodesic γ^* over γ can be isotoped into $\mathbb{S}^2$ without intersecting any other lift. This property is used in [Gabai et al. 2003].

In fact the solid cusp tubes associated with rank one parabolics are "unlinked" in a similar sense. For they are mutually disjoint, and any one of them can be shrunk towards a cusp (by using smaller and smaller horoballs) without bumping any other. Likewise the core curves in the solid tori obtained by Dehn filling on rank two cusps are unknotted, at least when the resulting manifold is close enough to the cusp.

On the other hand Otal defined unknottedness by a property suggested by the situation in $\mathbb{S}^3$: γ is *unknotted* if γ lies in an incompressible closed surface S properly embedded in the interior of $\mathcal{M}(G)$. If γ satisfies the latter definition, it also satisfies the former. For each lift S^* of S to $\mathbb{H}^3$ is a topological plane separating $\mathbb{H}^3$ and different lifts are mutually disjoint. Therefore γ^* can be moved slightly off S^* and then homotoped into $\mathbb{S}^2$ without hitting any lift of S. In particular if γ is a simple loop in the interior, homotopic to a point but not bounding an embedded disk in its complement, then γ is a knot in Otal's sense. Conversely if γ satisfies the former definition, it is a boundary component of a half-infinite cylinder extending out to an end. It is likely that γ is then isotopic to a simple loop on the boundary of a relative core. In this case it would also satisfy Otal's condition and the two definitions would be equivalent: for example, by intersecting the cylinder with a pleated surface exiting the end. In any case Otal's has been a fruitful definition. He proved the following beautiful theorem:

Unknottedness Theorem 6.6.6 [Otal 1995]. *Suppose there is a diffeomorphism* $\Phi : \mathcal{M}(G)^{\text{int}} \to S \times \mathbb{R}$. *There exists a constant* $0 < c = c(S)$ *such that any simple closed geodesic* $\gamma \subset \mathcal{M}(G)$ *of length* $<\ c$ *is isotopic to a curve* γ^* *contained in* $S^* = \Phi^{-1}(S \times \{0\})$. *Moreover the union of all simple closed geodesics* $\{\gamma_i\}$ *in* $\mathcal{M}(G)$ *of length* $<\ c$ *is isotopic to a union of simple loops in disjoint surfaces in* $\Phi^{-1}(S \times \mathbb{Z})$.

Therefore the collection of short curves is not only unknotted, but also unlinked. If $\mathcal{M}(G)$ is geometrically infinite, an infinite number of distinct short geodesics may exit one or both ends. (Such a sequence will exist in $\mathcal{M}(G)$ if and only if there is no positive lower bound for the length of the geodesics.)

Recently, Souto found the following generalization: There exists $\varepsilon_g > 0$ such that for all handlebodies $\mathcal{M}(G)$ of genus g, or compression bodies with incompressible surface of genus g, the set of geodesics in $\mathcal{M}(G)$ of length $< \varepsilon_g$ is unlinked.

At this point the prudent reader may think that Souto's result is vacuous because there is a positive lower bound for the length of a geodesic in a handlebody which depends only on the length of the shortest geodesic on its boundary. Such a belief is entirely wrong! Take a simple geodesic in the interior, and do higher and higher orders of Dehn surgery about it so that the manifolds nearly have rank two cusps. The length of the core geodesics will become arbitrarily small, independent of the length of the shortest geodesic on $\partial\mathcal{M}(G)$.

In addition, Souto (personal communication) showed that if $\mathcal{M}(G)$ is a closed manifold with a Heegaard splitting surface Σ_g of genus g and $\gamma \subset M(G)$ is a geodesic of length $< \varepsilon_g$, there is an embedded surface S isotopic to σ_g such that $\gamma \subset S$ (personal communication) . Short geodesics not only are simple, but at least in some cases and perhaps in all, are also unknotted and unlinked.

6-5. Consider a (finitely generated) quasifuchsian group G without parabolics. Given a compact submanifold K in the interior of $\mathcal{M}(G)$, show there exists $L > 0$ such that if γ is a closed geodesic of length not exceeding L, then γ lies in the complement of K and is parallel to a simple loop in a boundary component of $\mathcal{M}(G)$.

Prove that the manifold obtained by putting a hyperbolic structure on the result of drilling out a Myers curve Theorem 6.6.3 cannot appear as a geometric limit on $\partial\mathfrak{T}(G)$ [Soma 2003].

6-6. *Grafting.* This construction was used by Bill Goldman [Goldman 1987] to describe all real projective structures — those with fuchsian holonomy — over the deformation (Teichmüller) space Teich(R) of a closed Riemann surface R (see Exercise 6-8). Let G be a fuchsian group acting on the upper (UHP) and lower (LHP) half-planes and representing $R = \mathrm{LHP}/G$. Suppose the negative imaginary axis is the lift $\hat{c}$ of a simple closed geodesic $c \subset R$ of length L.

The element $g \in G$ determined by lifting c into $\hat{c}$ is $g : z \mapsto e^L z$. The quotient $T = (\mathbb{C} \setminus \{0\})/\langle g\rangle$ is a torus. There is a homeomorphic lift of c into T that we will also denote by c. Cut T along c to get a cylinder; the size of the cylinder — ratio of height to circumference — is determined by L, likewise cut R along c to get one or two subsurfaces. Attach the cylinder $T \setminus c$ to $R \setminus c$; that is pull open the cut of R and insert the cylinder $T \setminus c$ in the cut. This is possible because the ends of the cylinder have length L. We get a new Riemann surface R_c homeomorphic to R with a different conformal structure. Uniformize to get a representation $R_c = \mathrm{LHP}/G_c$.

We will now carry out this operation in $\mathbb{C}$: Slit $\mathbb{S}^2$ along $\hat{c}$ and denote the result by ${\mathbb{S}^2}_{\mathrm{cut}}$; the cut has $+$ and $-$ edges. Then slit LHP along $\hat{c}$ and denote the result by $\mathrm{LHP}_{\mathrm{cut}}$, it has corresponding $+$ and $-$ edges. Attach ${\mathbb{S}^2}_{\mathrm{cut}}$ to $\mathrm{LHP}_{\mathrm{cut}}$ by sewing the $-$ and $+$ edge of ${\mathbb{S}^2}_{\mathrm{cut}}$ to the $+$ and $-$ edge of $\mathrm{LHP}_{\mathrm{cut}}$. Likewise sew ${\mathbb{S}^2}_{\mathrm{cut}}$ equivariantly along all lifts of c to LHP. We get an abstract simply connected surface $\widehat{\mathbb{H}}$ lying over $\mathbb{S}^2$ and G acts on $\widehat{\mathbb{H}}$. The quotient $\widehat{\mathbb{H}}/G$ conformally equivalent to the

Riemann surface R_c we obtained by sewing into R a particular cylinder. Let π denote the projection of $\widehat{\mathbb{H}}$ onto $\mathbb{S}^2$. Under π, the action of G on $\widehat{\mathbb{H}}$ projects to the action of the original fuchsian group G on $\mathbb{S}^2$.

Now the abstract $\widehat{\mathbb{H}}/G$ is conformally equivalent to R_c so there is a conformal map $F : \text{LHP} \equiv \mathbb{H}^2 \to \widehat{\mathbb{H}}$ which conjugates G_c to G. The composition $f = \pi \circ F : \mathbb{H}^2 \to \mathbb{S}^2$ is a locally univalent meromorphic function; G is referred to as its monodromy group. This is an example of a *real projective structure* on R_c or on LHP with respect to G_c. It is called "real" because the monodromy group is a fuchsian group. The components of $f^{-1}(\mathbb{R})$ consist of mutually disjoint arcs which project to simple loops on R_c bounding an annulus (earlier called a cylinder) in the homotopy class of c.

The process of cutting a simple geodesic $c \subset R$ and inserting an annulus is called *grafting*. More particularly we have just done 2π-grafting along c as up in the universal covering we have wrapped the slit sphere once around.

A *multicurve* is a finite geodesic lamination λ, that is, it is a mutually disjoint set of simple geodesics on R. With the assignment of an integral weight $2\pi m_j$ to each component ℓ_j, where m_j is a positive integer, it becomes a measured lamination. Such measured laminations are denoted by $\mathcal{ML}_{\mathbb{Z}}$ or when considered projectively, $\mathcal{PML}_{\mathbb{Z}}$. The example we have just described c was assigned the weight 2π and we have accordingly grafted R. If the weight were instead $2\pi m$ we would have attached m copies of the cylinder $T \setminus c$ to $c \in R$ and m copies of the slit sphere to $\hat{c}$. Correspondingly, *integral grafting* can be effected by any element of $\mathcal{ML}_{\mathbb{Z}}$. In all cases the construction results in a structure $\widehat{\mathbb{H}}$ acted on by G, a new Riemann surface $R_\lambda =\text{LHP}/G_\lambda$, and a locally injective meromorphic function conjugating G_λ to G.

Once the lamination $\lambda \in \mathcal{ML}_{\mathbb{Z}}$ has been fixed, the integral grafting map can be interpreted as acting on the full deformation (Teichmüller) space, $\text{gr}_\lambda : \text{Teich}(R) \to \text{Teich}(R)$.

Theorem 6.6.7 [Tanigawa 1997]. *Integral grafting is a real analytic homeomorphism of* $\text{Teich}(R)$ *onto itself*.

Therefore for each $\lambda \in \mathcal{M}L_{\mathbb{Z}}$, there is a Riemann surface S_λ such that the λ-grafting on $S_\lambda = \mathbb{H}^2/G_\lambda$ is realized by a locally univalent function f_λ for $R = \mathbb{H}^2/G$. We will pursue this in Exercise 6-8.

Grafting can be defined for nonintegral weights on multicurves, and by continuity for any measured lamination (Λ, μ). As a consequence of [Goldman 1987], the more general graftings do not result in fuchsian holonomy. In general:

Theorem 6.6.8 [Scannell and Wolf 2002]. *Grafting is a homeomorphism of* $\text{Teich}(R)$ *onto itself*.

The construction of each convex core boundary can be interpreted as grafting on the component of $\partial\mathcal{M}(G)$ that it faces (see [Epstein et al. 2006] for a discussion).

6-7. *Constructing a cone manifold on an unknotted geodesic: Bromberg's construction.* This process is analogous to Exercise 6-6 but it is harder to implement. Let

$\mathcal{M}(G)$ be a quasifuchsian manifold whose bottom end $S = \partial_{\mathrm{bot}}\mathcal{M}(G)$ is a closed surface and whose top is geometrically infinite. Minsky's paper [2001] covers the case that there is a positive lower bound for the length of all geodesics. So assume there is a sequence of simple loops exiting the infinite end for which the lengths of the geodesic realizations $\{\gamma_i \in \mathcal{M}(G)\}$ are Otal-unknotted with lengths approaching 0.

Fix an index i. We can assume that $\gamma = \Phi(\sigma)$ lies in $S^* = \Phi(S \times \{t_i\})$. Assume for definiteness that γ does not divide S^*. A lift $\hat{\gamma}$ of γ to $\mathbb{H}^3$ determines a cyclic loxodromic subgroup $\langle g \rangle$ of G. The quotient $T = \mathbb{H}^3/\langle g \rangle$ is the interior of a solid torus, and we can designate its core loop again by γ.

The set $C = \Phi(\sigma \times [t, 1))$, $t_i \le t < 1$ is a half-open cylinder in the interior of $\mathcal{M}(G)$ which is bounded at one end by γ while its other end exits the top of $\mathcal{M}(G)$. The cylinder C lifts homeomorphically to give a half-open cylinder $\tilde{C}$ in the (open) solid torus T. Cut $\mathcal{M}(G)$ along C and T along $\widetilde{C}$. Isometrically glue $T \setminus \widetilde{C}$ to $\mathcal{M}(G) \setminus C$; $\widetilde{C}$ has two sides in T and likewise C in $\mathcal{M}(G)$. This will give a hyperbolic cone manifold $\mathcal{M}^*$ with cone axis γ and cone angle 4π, like a two-sheeted branched cover of the plane. Still $\mathcal{M}^*$ is homeomorphic to $\mathcal{M}(G)$, but now the hyperbolic structure is singular along γ.

Bromberg made the striking discovery that the hyperbolic cone manifold $\mathcal{M}^*$ is a quasifuchsian cone manifold: the bottom end is conformally the same as that of $\mathcal{M}(G)$ but the top end is now geometrically finite as well. An additional interesting fact that the subgroup $G_0 \cong \pi_1(\mathcal{M}(G) \setminus C)$ is a Schottky group. Even if G is degenerate or doubly degenerate, cutting along A has the profound effect on the limit set of totally disconnecting it. With Theorem 6.6.4 Bromberg then proved that the cone angle can be deformed to 2π without changing the conformal structures on the two ends. We end up with a geometrically finite, hyperbolic quasifuchsian manifold.

By applying the result to the sequence of simple loops $\{\sigma_i \subset S \times \{t_i\}\}$ for which the geodesic representatives of $\Phi(\sigma_i)$ shrink to zero, one can complete the proof of Bromberg's Theorem, stated on page 271.

6-8. *Projective structures.* Let $R = \mathrm{LHP}/G$ be a closed Riemann surface (the case with punctures is not so well understood), and G a fuchsian group. We recall that a Bers slice $\mathcal{B}(R)$ is the collection of *conformal* mappings $f : \mathrm{LHP} \to \mathbb{S}^2$ that have a *quasiconformal extension* to $\mathbb{S}^2$. The extension, which will also be denoted by f, induces an isomorphism θ to a quasifuchsian group by $f(g(z)) = \theta(g)(f(z))$, for all $z \in \mathbb{S}^2$, $g \in G$. The closure of the Bers slice is the closure of this space of conformal maps on LHP (modulo conjugation).

Add to the mix of conformal maps as follows. Take the much larger class of *locally injective meromorphic functions* $f : \mathrm{LHP} \to \mathbb{S}^2$ for which there exists a *homomorphism* $\varphi : G \to \mathrm{PSL}(2, \mathbb{C})$ satisfying $f(g(z)) = \varphi(g)(f(z))$ for all $z \in$LHP, $g \in G$. Locally injective meromorphic functions form the solution class of *schwarzian differential equations* on LHP over R (see Exercise 1-37),

$$S_\phi(f_\phi)(z) = \left(\frac{f''_\phi}{f'_\phi}\right)' - \frac{1}{2}\left(\frac{f''_\phi}{f'_\phi}\right)^2 = \phi(z), \qquad z \in \mathrm{LHP}.$$

Here ϕ is a lift of a holomorphic quadratic differential on R, namely it satisfies

$$\phi(g(z))g'(z)^2 = \phi(z), \ \forall z \in \text{LHP}, \ g \in G.$$

Solutions are uniquely determined by ϕ up to postcomposition with Möbius transformations so solutions can be appropriately normalized. The schwarzian derivative (the differential operator) is zero if and only if f is Möbius. A solution will in general map LHP *onto* $\mathbb{S}^2$.

On a *fixed* Riemann surface R there is the correspondence,

$$[f_\phi] \leftrightarrow \phi \leftrightarrow [\varphi_\phi],$$

where the brackets indicate PSL(2, $\mathbb{C}$) equivalence.

In current terminology, $f_\phi : R \to \mathbb{S}^2$ is called the *developing map*; it "unrolls" R over $\mathbb{S}^2$. The homomorphism φ associated with the differential ϕ or f_ϕ is called the *holonomy representation*. The *holonomy groups* or *monodromy groups* $\{\varphi(G)\}$ are in general not discrete or even finite presentable, but they are nonelementary. For a general introduction see for example [Gallo et al. 2000]. For another slant on projective structures relating to circle packings in $\mathbb{S}^2$, see [Kojima et al. 2006].

The collection of all projective structures on the *fixed* Riemann surface $R = \mathbb{H}^2/G$ (for definiteness we continue to think of $\mathbb{H}^2$ as LHP) is called the *extended Bers slice* $\mathcal{B}^*(R)$. It is parameterized by the $(3g-3)$-complex dimensional vector space of quadratic differentials on R. The extended slice $\mathcal{B}^*(R)$ is properly embedded in the representation variety $\mathfrak{R}(G)$ [Gallo et al. 2000]. We can ask about the components of its *discreteness locus*

$$\mathcal{B}^*_{\text{disc}}(R) = \{\varphi_\phi : \varphi_\phi \text{ is an isomorphism to a quasifuchsian group}\}.$$

This is an open set in $\mathcal{B}^*$. The component containing the basepoint (R, id) is the Bers slice.

Suppose we are given an element $(\lambda, \mu) \in \mathcal{ML}_\mathbb{Z}$, that is a multicurve λ with integral weights. There exists a uniquely determined Riemann surface $S_\lambda = \mathbb{H}^2/G_\lambda$ such that integral grafting on S_λ results in the Riemann surface R: The grafting map sends $\mathbb{H}^2$ over S_λ onto a simply connected surface $\widehat{\mathbb{H}}$ lying over $\mathbb{S}^2$ on which G_λ acts with $\widehat{\mathbb{H}}/G_\lambda$ conformally equivalent to R, as we saw in Exercise 6-6. The bottom line is that for each multicurve λ and assignment of integral weights giving an element of $\mathcal{ML}_\mathbb{Z}$, there is a projective structure f_ϕ on R such that the homomorphism φ corresponding to $f_\phi = \pi \circ F_\phi$ is an isomorphism $G \to G_\lambda$ while F_ϕ is a conformal map of $\mathbb{H}^2$ over R onto $\widehat{\mathbb{H}}$, conjugating the action of G to G_λ. The picture is filled out by the following important result, the second part of which has been studied and visualized by David Dumas:

Theorem 6.6.9 [Shiga and Tanigawa 1999; Dumas 2004]. *Each component of the discreteness locus of $\mathcal{B}^*(R)$ consists of quasifuchsian groups, and is biholomorphically equivalent to the Bers slice $\mathcal{B}(R)$. The component $\mathcal{B}_\lambda$ is indexed by its fuchsian center $\{c_\lambda = (R, \varphi_\phi)\}$. This point is determined by the quadratic differential ϕ on*

R with fuchsian monodromy group $\varphi_\phi(G) = G_\lambda$ with G_λ determined by that surface $R_\lambda = \mathbb{H}^2/G_\lambda$ on which λ-grafting yields R.

Each component of $\mathcal{B}^*(R) \cap \mathfrak{T}(G)$ is a "generalized Bers slice" consisting of certain deformations of a projective structure on R.

The theorem applies to finite area surfaces R more generally. When there are punctures the projective structures must be taken so that $\varphi(g)$ is parabolic whenever $g \in G$ is so.

In the case of a once-punctured torus, the 1-dimensional space of quadratic differentials for which the holonomy map preserves the parabolics, can be given explicitly. The first computation and visualization of the resulting Bers slice was carried out by the Japan team (Komori, Sugawa, Wada, Yamashita) in terms of complex probabilities. The result was dramatic. More recently they have computed the extended Bers slice. Using his own software, David Dumas has provided slices in addition for the hexagonal torus and more generally has broadened their explorations. See [Komori and Sugawa 2004; Wada 2006; Dumas 2004]. The frontispiece consists of one of Dumas' pictures which shows the Bers slice in the archipelago of components of $\mathcal{B}^*_{\text{disc}}$ with their fuchsian centers indicated.

We can also consider the totality of all projective structures on *all* Riemann surfaces in the deformation space $\text{Teich}(R) \equiv \text{Teich}(G)$, $R = \mathbb{H}^2/G$. This is given by the $(6g-6)$-complex dimensional bundle of quadratic differentials $\mathfrak{Q}(G)$ over $\text{Teich}(G)$ (equivalently, we may write $\mathfrak{Q}(R)$ over $\text{Teich}(R)$). The solution of the schwarzian equation for ϕ on $R' = \mathbb{H}^2/G'$ gives rise to the holonomy representation $\varphi_\phi \in \mathfrak{R}(G)$ onto the holonomy group $\varphi_\phi(G)$ (as usual, quotienting out by conjugations). According to [Gallo et al. 2000] the totality of monodromy groups $\{\varphi_\phi(G)\}$, $\phi \in \mathfrak{Q}(G)$ comprise the component $\mathfrak{R}_+(G)$ of $\mathfrak{R}(G)$ consisting of those representations that lift to $\mathrm{SL}(2,\mathbb{C})$: $\varphi(G)$ lifts if each generator can be assigned a matrix so that the designated matrices satisfy the surface relation satisfied by $\pi_1(R)$; the lifted group need not be isomorphic of $\varphi(G)$. Representations that so lift comprise one of the two components of $\mathfrak{R}(G)$.

The surjective holomorphic map $\mathrm{Hol} : \mathfrak{Q}(G) \to \mathfrak{R}_+(G)$ is a local homeomorphism, but it is not a covering mapping — closed arcs do not necessarily lift in their entirety [Hejhal 1975]. In Hejhal's examples this occurs because continuation over the path leads to a pinching of the underlying Riemann surfaces, which occurs before continuation is complete.

In the present context, the *discreteness locus* is defined as the closed set

$$\mathfrak{Q}_{\text{disc}}(G) = \{\phi \in \mathfrak{Q}(G) : \varphi_\phi \text{ is an isomorphism to a discrete group}\}.$$

It is analogous to the discreteness locus $\mathfrak{R}_{\text{disc}}(G)$ considered in §5.2.2. Its interior $\text{Int}(\mathfrak{Q}_{\text{disc}}(G))$ consists of quasifuchsian groups. Hejhal [1975] proved that Hol is injective on the components.

The extended Bers slice directly involves (in our setup) only LHP. The action of nonzero integral grafting sends the extended Bers slice based on $R = \mathrm{LHP}/G$ to the extended slice based on the surface $S_\lambda = \mathrm{LHP}/G_\lambda$.

6-9. *Self-bumping.* Suppose X is a component of $\mathfrak{R}_{\mathrm{disc}}$. We ask, when does X bump itself? That is, when is there a point $\zeta \in \partial X$ such that for all small enough neighborhoods $U \subset \mathfrak{R}(G)$ of ζ, $U \cap \mathrm{Int}(X)$ is not connected.

Making use of the technique of [Anderson and Canary 1996a], the prototypical case of a fuchsian group G representing a closed surface $R = \mathbb{H}^2/G$ was analyzed by Curt McMullen. (For a more general case see Theorem 5.3.2.) Because of its importance in future developments and its intrinsic interest, we will outline the argument.

Theorem 6.6.10 [McMullen 1998, Appendix]. *There exists a cusp ζ on the boundary of quasifuchsian space $\mathfrak{T}(G)$ such that for all small neighborhoods $U \subset \mathfrak{R}(G)$ of ζ, $U \cap \mathfrak{T}(G)$ is not connected.*

The closure $\overline{\mathfrak{T}(G)} \subset \mathfrak{R}(G)$ is not a manifold with boundary.

The theorem does *not* say that the closure $\overline{\mathfrak{T}(G)}$ is not locally connected. However, Bromberg [2006] has proved that $\overline{\mathfrak{T}(G)}$ is *not* locally connected in the once punctured torus case — for this case the boundary of a Bers slice is known to be locally connected (Minsky). Bromberg's clever proof takes advantage of the explicitness of once-punctured torus groups. This is the only deformation space local connectedness is known to hold.

The first construction is called wrapping. It results in an immersion of R which is not homotopic to an embedding. The construction originates in [Anderson and Canary 1996a] — see Exercise 5-13.

We will start by constructing the target manifold. Let d be a simple geodesic on say $\partial_{\mathrm{bot}}\mathcal{M}(G)$. Inside $\mathcal{M}(G) \cong R \times [0, 1]$, set $\delta = d \times \{\frac{1}{2}\}$. The manifold $\mathcal{M}(G) \setminus \{\delta\}$ is represented by a geometrically finite $\mathcal{M}(H)$. For an explicit construction see Exercise 4-19. We can take $\partial\mathcal{M}(H)$ to be conformally equivalent to the bottom and top components of $\partial\mathcal{M}(G)$.

Suppose for purposes of explaining the wrapping, $R = \mathbb{H}^2/G$ is a closed surface of genus two. Let $d \subset R$ be a simple loop that divides R into two subsurfaces of genus one. Let $a_1, b_2; a_2, b_2$ be simple loops about each of the two handles so that,

$$\pi_1(R) \cong \langle a_1, b_2, a_2, b_2 \big| [a_1, b_1] = [a_2, b_2] \rangle.$$

Taking a thin torus about δ with simple loops d' parallel to d and meridian c, we find that H has the presentation

$$H \cong \langle a_1, b_2, a_2, b_2, c, d : d = [a_1, b_1] = [a_2, b_2],\ [c, d] = 1 \rangle.$$

"Wrap" a closed surface homeomorphic to R about δ as follows. Start with the surface $S_0 = R \times \{3/4\}$. Take a solid torus in $\mathcal{M}(G)$ with core curve δ bounded by a torus T_0 containing $d \times 3/4$ but otherwise disjoint from S_0. Slit S_0 and T_0 along $d \times 3/4$. Join T_0 to S_0 by cross identifying over the slits to obtain an immersion

$f: R \to f(R) = S$. The immersed surface S wraps once around δ. The subgroup $\pi_1(S) \subset H$ is generated by $(a_1, b_1, ca_2c^{-1}, cb_2c^{-1})$. The immersion is not homotopic in $\mathcal{M}(G) \setminus \{\delta\} \cong \mathcal{M}(H)$ to an embedding, for $\pi_1(S)$ is not conjugate in H to the fundamental group of either component of $\partial\mathcal{M}(H)$.

We digress to make two remarks.

(i) The wrapping of R about δ can be done any integral number of times, and can be done for any collection of mutually disjoint simple loops $\{d\}$ on R.

(ii) There are intrinsic restrictions on wrapping, which are hidden under our assumption that the top and bottom of $\mathcal{M}(G)$ are closed surfaces. Suppose more generally they are finitely punctured surfaces. First of all, $d_\pm$ on the top and bottom of $\mathcal{M}(G)$ cannot be homotoped on $\partial\mathcal{M}(G)$ to a puncture. Secondly, if H is to exist, no loop which has nonzero geometric intersection with one of $d_\pm$ can determine a parabolic transformation in H. Thus if $\mathcal{M}(G)$ were, for example, a maximal cusp group on $\mathfrak{T}(G)$, the construction would be impossible.

Let $\varphi : G \to G^* \subset H$ be the isomorphism induced by the immersion $f: R \to S$. We deduce

(i) $\mathcal{M}(G^*)$ covers $\mathcal{M}(H)$.
(ii) $\mathrm{Int}\mathcal{M}(G^*) \cong R \times (0, 1)$.
(iii) No component of $\partial\mathcal{M}(G^*)$ corresponds to a component of $\partial\mathcal{M}(H)$ (the fundamental groups are not conjugate).
(iv) $\varphi : G \to G^* \in \partial\mathfrak{T}(G)$.

One component $\partial\mathcal{M}(G^*)$ is a closed surface homeomorphic to R and the other is the union of two surfaces sharing a parabolic resulting from pinching.

The point (φ, G^*) is a cusp on the boundary of some Bers slice. As such it is the limit of quasifuchsian groups in the slice. Specifically, there is a sequence of isomorphisms $\{\varphi_n : G \to G_n\} \in \mathfrak{T}(G)$ which converge algebraically and geometrically to $\varphi : G \to G^*$. In particular, $\lim \Omega(G_n) = \Omega(G^*)$.

The second construction required is the application of $(1, n)$ Dehn surgery on the cusp of $\mathcal{M}(H)$ where 1 corresponds to the meridian c, and n to the longitude d. Take a cusp torus, remove its interior, and replace it by a solid torus so that $\gamma = cd^n$ becomes the meridian, that is, $\gamma = cd^n \sim 1$. For all large n there results a hyperbolic manifold $\mathcal{M}(H_n)$ which is a quasifuchsian manifold homeomorphic to $\mathcal{M}(G)$. Carried along is the immersion

$$F_n : R \xrightarrow{f} S \subset \mathcal{M}(H) \xrightarrow{\text{inclusion}} \mathcal{M}(H_n),$$

which is homotopic to an embedding now that δ is no longer there. It induces the algebraically converging isomorphisms

$$\{\rho_n : G \to H_n\} \xrightarrow{n\to\infty} \varphi : G \to G^*.$$

On the other hand, $\{\mathcal{M}(H_n)\}$ converges geometrically back to $\mathcal{M}(H)$ (Exercise 5-11). According to Theorem 4.5.4, $\lim \Omega(H_n) = \Omega(H)$.

To complete the argument we must draw on the theory of projective structures $\mathfrak{Q}(R)$ as outlined in Exercise 6-8.

The representation $\varphi : G \to G^*$ in $\partial\mathfrak{T}(G)$ is the holonomy representation of some $\phi \in \mathfrak{Q}_{\text{disc}}(G)$. Let U be any small enough neighborhood of ϕ in $\mathfrak{Q}(G)$. For U small, the holonomy representation

$$\text{Hol} : U \mapsto V = \text{Hol}(U) \subset \mathfrak{R}(G)$$

is a homeomorphism onto a neighborhood V of $(\varphi, G^*) \subset \mathfrak{R}(G)$.

Set $U_d = U \cap \mathfrak{Q}_{\text{disc}}(G)$. There exists $\{\phi_n\} \subset U_d$ such that its holonomy is $\varphi_n : G \to G_n$. There also exists $\{\phi'_n\} \subset U_d$ with holonomy $\rho_n : G \to H_n$. Both sequences converge to ϕ. Now $\lim \Omega(G_n) = \Omega(G^*)$ while $\lim \Omega(H_n) = \Omega(H)$. Because $\Omega(G^*)$ and $\Omega(H)$ have no component in common it follows from looking at the corresponding developing mappings that $\{\phi_n\}$ and $\{\phi'_n\}$ cannot lie in the same component of U_d. Therefore $\{(\varphi_n, G_n)\}$ and $\{(\rho_n, H_n)\}$ do not lie in the same component of $\text{Hol}(U_d) = V \cap \mathfrak{T}(G)$. Yet both sequences converge to $(\varphi, G^*) \in \partial\mathfrak{T}(G)$.

It follows as a consequence of Theorem 5.10.12 that the closure $\overline{\mathfrak{T}(G)}$ is not a manifold.

6-10. Expanding on Theorem 6.6.10, Kentaro Ito [2000a; 2000b] made a detailed study of the situation and proved (compare with Theorem 6.6.9):

Theorem 6.6.11. *The components of* $\text{Int}(\mathfrak{Q}_{\text{disc}}(G))$ *are in one-to-one correspondence with the elements of* $\mathcal{ML}_{\mathbb{Z}}$.

Denote the component of the interior corresponding to $\lambda \in \mathcal{ML}_{\mathbb{Z}}$ *by* $\mathfrak{Q}_\lambda$; $\mathfrak{Q}_0 = \mathfrak{T}(G)$. *The component* $\mathfrak{Q}_\lambda$ *has a uniquely determined fuchsian center* c_λ: *there is a surface* $S_\lambda = \mathbb{H}^2/G_\lambda$ *on which* λ*-grafting determines a projective structure on* R *with monodromy group* G_λ. $\mathfrak{Q}_\lambda$ *is biholomorphically equivalent to quasifuchsian space* $\mathfrak{T}(G) = \mathfrak{Q}_0$.

Choose any $\lambda, \mu \in \mathcal{ML}_{\mathbb{Z}}$.

(i) *While* Hol : $\mathfrak{Q}_\lambda \to \mathfrak{R}(G)$ *is injective* [Hejhal 1975], *it is not injective on the closure* $\overline{\mathfrak{Q}}_\lambda$ *unless* $\lambda = 0$.
(ii) *For the closures in* $\mathfrak{Q}(G)$, $\overline{\mathfrak{Q}}_\lambda \cap \overline{\mathfrak{Q}}_\mu \neq \varnothing$; *in particular* $\overline{\mathfrak{Q}}_\lambda \cap \overline{\mathfrak{Q}}_0 \neq \varnothing$.
(iii) *There exists* $\zeta \in \partial\mathfrak{Q}_0 \cap \partial\mathfrak{Q}_\lambda$ *such that for all small neighborhoods* U *of* ζ, $U \cap \mathfrak{Q}_\lambda$ *is not connected*; $\overline{\mathfrak{Q}}_\lambda$ *is not a manifold.*
(iv) *The closed set* $\mathfrak{Q}_{\text{disc}}(G)$ *is connected.*

Another way of identifying $\mathfrak{Q}_\lambda$ is that it contains the fuchsian center c_λ.

6-11. *Meromorphic functions and laminations.* This construction is due to Thurston; a proof appears in [Kamishima and Tan 1992]. Suppose $f : \mathbb{D} \to \mathbb{S}^2$ is a locally injective meromorphic function in the unit disk $\mathbb{D} \equiv \mathbb{H}^2$. Consider round disks $\{D \subset \mathbb{S}^2\}$ with the property that there is a single valued branch of $f^{-1} : D \to \mathbb{D}$. We may assume all such disks $\mathcal{D} = \{D\}$ are maximal in the sense none is contained in a larger disk on which f^{-1} has a branch.

Consider the set $\mathcal{U} = \{U = f^{-1}(D) : D \in \mathcal{D}\}$ and set $U^\infty = \overline{U} \cap \partial\mathbb{D}$; each U^∞ contains at least two points. Construct the hyperbolic convex hull $\mathcal{C}(U^\infty)$ in $\mathbb{H}^2$. The following properties hold:

(i) Corresponding to each point $z \in \mathbb{D}$ is a unique element $U_z \in \mathcal{U}$ such that $z \in \mathcal{C}(U^\infty)$.
(ii) Two hulls $\mathcal{C}(U_1^\infty), \mathcal{C}(U_2^\infty)$ are either disjoint, or they share a common edge and/or vertex on $\partial\mathbb{D}$.
(iii) The collection of hulls $\{\mathcal{C}(U^\infty)\}$ covers $\mathbb{D}$ without overlapping interiors.
(iv) If $f \circ g(z) = \varphi(g) \circ f(z)$ for a homomorphism $\varphi : G \to \mathrm{PSL}(2, \mathbb{C})$, for all elements g of a fuchsian group G and all $z \in \mathbb{D}$, then the action of G permutes the elements of $\{\mathcal{C}(U^\infty)\}$ while the elements of $\varphi(G)$ permute the maximal disks in $\mathcal{D}$.

In this analysis, it is helpful to examine for a given $z \in \mathbb{D}$, the set $W_z = \bigcup_{z \in U} U$, the union of those elements of $\mathcal{U}$ that contain z. Then set $W_z{}^\infty = \overline{W}_z \cap \partial\mathbb{D}$. The image set $f(W_z^\infty) \subset \mathbb{S}^2$ is well defined and we can pass to its convex hull $\mathcal{C}(fW_z^\infty)$, now taken with respect to $\mathbb{H}^3$. There is a "closest" point $r(f(z)) \in \mathcal{C}(fW_z^\infty)$ to $f(z)$, where r denotes the nearest point retraction. Construct the hyperbolic plane $P_{f(z)}$ which is orthogonal to the segment $[f(z), r(f(z))]$ at the point $r(f(z))$. The boundary of $P_{f(z)}$ on $\partial\mathbb{S}^2$ is a circle, and one of the disks $D_{f(z)}$ that it bounds on $\mathbb{S}^2$ contains $f(z)$. In fact $D_{f(z)} \in \mathcal{D}$ and $f^{-1}(D_{f(z)}) \in \mathcal{U}$. The map $\Psi : \mathcal{C}(U_z^\infty) \subset \mathbb{H}^2 \to \mathcal{C}(fW_z^\infty)$ sending u to $r(f(u))$ is an isometry; more generally the map $z \in \mathbb{D} \to r(f(z)) \in \mathbb{H}^3$ determines an isometry to a pleated surface in $\mathbb{H}^3$.

The set of all edges of the convex hulls in $\mathbb{H}^2$ form a geodesic lamination Λ. It will be invariant under the group G, if there is one. There is a naturally determined bending measure on this lamination. Namely if for $U_1, U_2 \in \mathcal{U}$, $U_1 \cap U_2$ intersect with exterior angle $0 < \alpha < \pi$, then so do the image disks in $\mathcal{D}$, since f is locally a conformal mapping. The corresponding $\mathcal{C}(U_1^\infty)$, $\mathcal{C}(U_2^\infty)$ share an edge. We assign the bending angle α to that. However the geodesics in Λ are unlikely to be isolated and then a process akin to Riemann integration is used to obtain the bending measure.

Thurston's insight was that this construction, disseminated to the world by Bill Goldman and proved in [Kamishima and Tan 1992], results in a coordinate system for the projective space $\mathfrak{Q}(G)$, where the fuchsian group G can represent either a closed or a finite area surface:

Theorem 6.6.12. *The construction described above results in a homeomorphism* $\Theta : \mathfrak{Q}(G) \to \mathfrak{T}(G) \times \mathcal{ML}(G)$.

That the map is surjective is a consequence of the fact that grafting on a Riemann surface R' gives rise to a projective structure on another Riemann surface S', as described in Exercise 6-6.

Theorem 6.6.12 shows that corresponding to each Riemann surface $R' \in \mathrm{Teich}(R)$ and projective structure on it is a uniquely determined Riemann surface S_λ and measured lamination λ on S_λ with the following property: Grafting on S_λ determines the given structure on R'. However we can no longer be restricted to integral grafting. That is, $\mathfrak{Q}(R) \cong \mathrm{Teich}(R) \times \mathcal{ML}(R)$. These are called the *Thurston coordinates*. Using the Thurston coordinates, David Dumas [2004] has proved that one can compactify the space so that $\overline{\mathfrak{Q}(R)} \cong \mathcal{PML}(R) \times \mathcal{PML}(R)$.

6-12. *Hyperbolic manifolds with corners.* This exercise was inspired by [Otal 1998, §§7,8], to which the reader is referred for more detail and for application to the proof of the Hyperbolization Theorem.

Start with a compact, acylindrical manifold $M = \mathcal{M}(X)$ with nonempty boundary. Let $\{R_i\}$ be a decomposition of ∂M into compact, (connected) subsurfaces with mutually disjoint interiors. That is, each R_i is a compact bordered surface, or an entire boundary component of M. But we require that R_i is not a topological annulus or disk.

Associated with the decomposition is the decomposition graph $\mathcal{G}$: Each vertex corresponds to an R_i, and two vertices are joined by an edge if the subsurfaces are adjacent along a common border. We will call the graph an *admissible decomposition graph* and the corresponding $\{R_i\}$ an *admissible decomposition*, if we can give each R_i a label $+$ or $-$ in such a way that if R_i and R_j share a boundary component, they have different labels. Thus every simple loop in $\mathcal{G}$ is composed of an even number of edges.

Suppose then we have an admissible decomposition of ∂M. Denote the subsurfaces with the label $+$ by $(\Sigma_1, \Sigma_2, \ldots)$, and the subsurfaces with the $-$ label—the complementary subsurfaces—by $(\Sigma'_1, \Sigma'_2, \ldots)$.

Carry out the same "reflection" process as in Exercise 6-2 with respect to the b subsurfaces $\{\Sigma_i\}$. For example, if Σ_i is adjacent to Σ'_j along γ, after reflection J in Σ_i, the reflected surface $J(\Sigma'_j)$ is attached to Σ'_j along γ. There results an orientable, 2^b-sheeted covering $\widehat{M_{\Sigma'}}$ whose boundary is the union of closed orientable surfaces. Each of its boundary components is the union of lifts of elements of $\{\Sigma'_i\}$ and each element of $\{\Sigma'_i\}$ is represented 2^b times in the boundary. Again $\widehat{M}_{\Sigma'}$ is an orientable, irreducible, acylindrical, atoroidal, Haken manifold.

And again there is a group of automorphisms G of order 2^b acting on $\widehat{M}_{\Sigma'}$ generated as above by the "reflections" in the elements of $\{\Sigma_i\}$.

Realize $\widehat{M}_{\Sigma'}$ as a hyperbolic manifold $\mathcal{M}(Y)$; we may assume that its boundary is totally geodesic. The "reflections" we introduced in the surfaces $\{\Sigma_i\}$ now become orientation reversing isometries of $\mathcal{M}(Y)$. A "reflection" in Σ_i corresponds to an isometry of $\mathcal{M}(Y)$ which pointwise fixes a totally geodesic subsurface, which we will again label Σ_i. The boundary of each Σ_i is contained in $\partial\mathcal{M}(Y)$, which Σ_i meets orthogonally.

The quotient $\mathcal{M}(Z) = \mathcal{M}(Y)/G$ is homeomorphic to the result of cutting $\mathcal{M}(Y)$ along the surfaces $\{\Sigma_i\}$ hanging orthogonally from its boundary. Thus it is homeomorphic to the original M. In particular $\partial\mathcal{M}(Z)$ is the union of compact subsurfaces which we can again label as $\{\Sigma_i\}$ and $\{\Sigma'_i\}$.

Our investigations can be summarized as follows.

Hyperbolic Manifolds with Corners 6.6.13. *Given M as above, there is a uniquely determined kleinian group Z with the following properties*:

(i) *$\mathcal{M}(Z)$ is compact and $\partial\mathcal{M}(Z)$ is the union of two systems of subsurfaces $\{\Sigma_i\}$, and $\{\Sigma_i'\}$. None of the subsurfaces are simply or doubly connected. The totality of interiors are mutually disjoint. The closures of the members of each system are mutually disjoint.*

(ii) *The ordinary set $\Omega(Z)$ is the union of round disks $\{D_i\}$. Either $D_i \cap D_j = \varnothing$ for all $j \neq i$, or D_i is orthogonal to D_j for some $j \neq i$. In the former case, D_i covers a boundary component of $\mathcal{M}(Z)$ which is a member of one of the systems. In the latter case, D_i and D_j contain lifts of subsurfaces of different systems, and the stabilizer of $D_i \cap D_j$ is the cyclic group determined by a common boundary component.*

Note the symmetry in properties between the two systems of subsurfaces. An interesting special case is when the elements of the systems comprise a pants decomposition.

6-13. *Residual finiteness.* A group G is said to be *residually finite* if for any $g \neq \mathrm{id} \in G$, there exists a subgroup H of finite index such that $g \notin H$. A subgroup of a residually finite group is also residually finite [Hempel 1976].

Residual finiteness is known to hold for surface groups. Even more strongly, suppose S is a surface of possibly infinite topological type and $G \subset \pi_1(S)$ is a finitely generated proper subgroup. Choose $g \in \pi_1(S) \setminus G$. There exists a finite sheeted covering surface S^* of S such that $G \subset \pi_1(S^*)$ with injective inclusion $G \hookrightarrow \pi_1(S^*)$, but that $g \notin \pi_1(S^*)$; see [Scott 1978].

On the other hand, for 3-manifolds, Hempel [1987] proved that the fundamental group of a (compact) Haken manifold is residually finite. In fact every finitely generated matrix group is residually finite. It follows that the fundamental group of every geometric 3-manifold proclaimed in the Geometrization Conjecture/Theorem is residually finite [Thurston 1982b, Theorem 3.3].

6-14. *Infinitely generated kleinian groups.* Riemann surfaces of infinite genus and/or an infinite number of ends can be represented by fuchsian groups. Each such fuchsian group has in turn a quasiconformal deformation space. The elementary combination theorems can be used to paste together such groups over round disks or horodisks to construct a range of infinitely generated (non-finitely-generated) groups, just as in the finitely generated case. Likewise infinitely generated Schottky groups can be constructed.

Is there a classification of infinitely generated kleinian groups?

Here is an interesting example of Bromberg and Souto (private communication).

Start with a closed hyperbolic manifold M with a nondividing incompressible surface $S \subset M$. Set $N = M \setminus S$. Assume that N does not have the form $S \times (0, 1)$. Then $\pi_1(S)$ is a proper subgroup of $\pi_1(N)$. We may choose the hyperbolic structure on N so as to have totally geodesic boundary components $S_-, S_+ \cong S$ (it will not be fuchsian).

Set $N_0 = N$. Form the hyperbolic manifold N_1 by reflecting N_0 across S_+. In other words take two copies of N_0 and glue the top boundary component S_+ of N_0 to the bottom S_- of the copy. Correspondingly we can form N_{-1}. Let N_k denote the hyperbolic manifold formed from N_0 by successively gluing together the $2k+1$ copies of N_0: glue to N_0 k copies in the positive and in the negative direction.

N_k is a hyperbolic manifold whose two boundary components are conformally equivalent to S_-, S_+. We have that $\pi_1(N_k)$ is a proper subgroup of $\pi_1(N_{k+1})$. Normalize the representations $N_k = \mathcal{M}(G_k)$ so that $G_k \subset G_{k+1}$, $k = 1, 2, \ldots$. Then $\{G_k\}$ converges algebraically and geometrically to a infinitely generated group H with $\mathcal{M}(H) \ncong S \times \mathbb{R}$. Its limit set is all $\mathbb{S}^2$.

Bromberg and Souto show as a consequence of the finite area of the boundary components of the approximates $\{\mathcal{M}(G_k)\}$ that $\mathcal{M}(H)$ is quasiconformally rigid: any quasiconformal conjugation of H to another group is Möbius. They then show that if M^* is another hyperbolic manifold that is homeomorphic to $M(H)$, there is a bilipschitz map between them. By quasiconformal rigidity the two manifolds are in fact isometric.

More generally, they conjecture that any manifold $\mathcal{M}(H)$ which can be exhausted by tame manifolds whose boundaries have uniformly bounded areas is either rigid or it has a tame end. Soma's work (Theorem 5.9.2) gives a multitude of infinitely generated examples.

Agol conjectured that any irreducible, atoroidal, orientable 3-manifold with infinitely generated fundamental group is hyperbolic provided the covering corresponding to any finitely generated subgroup is tame.

7

Line geometry

This chapter describes an elegant method that makes it easier to quantitatively analyze geometric situations in hyperbolic space that involve lines and planes.

7.1 Half-rotations

We will identify each line $\ell \in \mathbb{H}^3$ with the *half-rotation* about ℓ, that is, the elliptic transformation of order 2 having ℓ as its axis.

The Cayley–Hamilton identity satisfied by normalized matrices A is

$$A + A^{-1} = \tau_A I, \quad \text{or} \quad A^2 = \tau_A A - I. \tag{7.1}$$

As always, τ_A is the trace and I is the 2×2 identity matrix. In particular, $\tau_{(A^2)} = (\tau_A)^2 - 2$.

Lemma 7.1.1. *Half-rotations correspond to normalized matrices A of trace $\tau_A = 0$ (eigenvalues $\pm i$), or equivalently, matrices A that satisfy*

$$A^2 = -I. \tag{7.2}$$

Consider the half-rotation $M = \left(\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right)$, with $-a^2 - bc = 1$ and $c \neq 0$. Its fixed points are $P = (a+i)/c$ and $Q = (a-i)/c$. Thus we can write

$$M = \frac{i}{(P-Q)} \begin{pmatrix} P+Q & -2PQ \\ 2 & -(P+Q) \end{pmatrix}. \tag{7.3}$$

If P or Q is ∞, take the corresponding limit of the expression.

The determinant of the matrix, $-(P-Q)^2$, does not have a uniquely determined root. Correspondingly, there is no way to distinguish between the two fixed points P, Q.

We are reminded of the useful Lemma 2.1.3:

Lemma 7.1.2. *A Möbius transformation T interchanges two distinct points x, y on $\mathbb{S}^2$ if and only if it is a half-rotation. If so, T is the half-rotation about a line ℓ orthogonal to the line τ between x and y.*

Proof. We may assume τ is the vertical axis in the upper half-space model. The map T maps τ onto itself, interchanging its endpoints. Therefore T has a fixed point $p \in \tau$. T^2 fixes the two endpoints and also fixes p so it must be the identity. Let P be the plane through p orthogonal to τ. T also maps P onto itself, but interchanges its two sides in $\mathbb{H}^3$. Therefore the rotation axis ℓ of T lies in P and is necessarily orthogonal to τ at p.

Conversely, suppose T is the half-rotation about ℓ. Taking ℓ to be the vertical conjugates T to the map $z \mapsto -z$. It interchanges the opposite points on every concentric circle about $z = 0$. $\square$

7.2 The Lie product

The *Lie product* φ of two nonsingular 2×2 matrices A, B is defined as

$$\varphi = AB - BA =: \{A, B\}.$$

Interchanging A and B changes φ to $-\varphi$ but leave the corresponding Möbius transformation unchanged. Likewise the Möbius transformation corresponding to the matrix φ is independent of the sign chosen for A or B.

We have the relations

$$U\{A, B\}U^{-1} = \{UAU^{-1}, UBU^{-1}\} \tag{7.4}$$

and

$$\det \varphi = 2 - \operatorname{tr}(ABA^{-1}B^{-1}). \tag{7.5}$$

The second one follows from the identity $\varphi = (ABA^{-1}B^{-1} - I)BA$ and the formula $\det(X - I) = 2 - \tau_X$. If $\det \varphi \neq 0$, since φ has zero trace,

$$\varphi^2 = -(\det \varphi)^2 I.$$

Of course, interpreted as a Möbius transformation, $\varphi^2 = \mathrm{id}$.

In preparation for the next result, note that two distinct lines in $\mathbb{H}^3$ not having a common endpoint always have a common perpendicular. This can be seen in the upper half-space model by taking one of the lines as the vertical axis. If the two lines intersect in $\mathbb{H}^3$, the common perpendicular is the line through the point of intersection and orthogonal to the plane containing the two lines.

Proposition 7.2.1. *The Lie product $\varphi = \{A, B\}$ has the following properties*:

(*1*) $\varphi = 0$ *if and only if either*

- *A and/or B is kI for some $k \in \mathbb{C}$, or*
- *A and B correspond to transformations with the same set of fixed points on $\partial\mathbb{H}^3$.*

More generally, φ *is singular* ($\det\varphi = 0$) *if and only if* $\operatorname{tr}(ABA^{-1}B^{-1}) = +2$. *In other words,* φ *is singular if and only if* A *and* B *have at least one fixed point in common on* $\partial\mathbb{H}^3$.

(2) $$\varphi A^{-1} = A\varphi, \quad \varphi B^{-1} = B\varphi. \tag{7.6}$$

(3) *If* φ *is nonsingular, it corresponds to a half-rotation. Its axis is:*

- *The common perpendicular to the axes of* A *and* B, *if neither is parabolic.*
- *The line between the fixed points if both* A *and* B *are parabolic, or, if only one is parabolic, the line from the parabolic fixed point which is orthogonal to the axis of the other.*
- *The line between the diametrically opposite points* $\mathcal{C} \pm i/c$ *on the isometric circle of* φ *with center* $\mathcal{C}$ *and radius* $|c|^{-1}$, *where* c *is the lower left term in the normalized matrix for* φ.

(4) *If* $\tau_A \neq \pm 2$, *then* $\psi = A - A^{-1}$ *is the half-rotation about the axis of* A.

(5) *Suppose* A *and* B *are half-rotations. Then* AB *has the same axis as* φ. *If the axes of* A *and* B *intersect at a point* $x \in \mathbb{H}^3$, *then* AB *is necessarily elliptic and fixes* x *as well. If all three* A, B, AB *are half-rotations, then the three rotation axes have a common point of intersection in* $\mathbb{H}^3$ *and are mutually orthogonal there.*

(6) *If only* B *is a half-rotation, its axis is orthogonal to the axis of* A, (*or ends at the fixed point of* A *if* A *is parabolic*), *if and only if* AB *is a half-rotation. If this is the case, the axis of* AB *is also orthogonal to that of* A (*ends at the fixed point of* A *if* A *is parabolic*).

(7) *The axes of* A, B, C *have a common perpendicular if and only if,*

$$\tau_{ABC} = \tau_{CBA}. \tag{7.7}$$

If A, B, C *are all half-rotations, the condition for a common perpendicular becomes*

$$\tau_{ABC} = 0. \tag{7.8}$$

(8) *If* $aA + bB + cC = 0$ *for nonzero scalars* a, b, c *and matrices representing nonparabolic elements, then the axes of the transformations corresponding to* A, B, *and* C *have a common perpendicular.*

In the special case that A and B preserve the upper half-plane UHP, they also preserve the vertical half-plane H in $\mathbb{H}^3$ based on $\mathbb{R}$. Their axes lie in H. If their axes intersect, then the axis of φ is orthogonal to H, and passes through the point of intersection; φ itself also preserves UHP and H. If instead the axes of A and B are disjoint, then the axis of φ lies in H as well and φ interchanges the upper and lower half-planes.

Example 7.2.2. Let $\mathcal{P}$ denote a regular hyperbolic octagon in $\mathbb{H}^2$ with vertex angles $\pi/4$. Going around $\partial\mathcal{P}$ in its positive direction, label its edges a, b, c, d, a^{-1}, b^{-1}, c^{-1}, d^{-1}. Let A denote the Möbius transformation that maps a to a^{-1}, sending $\mathcal{P}$

to the right side of a^{-1}, and similarly find transformations B, C, D. The group $G = \langle A, B, C, D\rangle$ is fuchsian and represents a genus 2 surface. Its generators satisfy the relation $ABCDA^{-1}B^{-1}C^{-1}D^{-1} = \mathrm{id}$, or $ABCD = DCBA$. Therefore ABC is conjugate to CBA and BCD to DCB. By property (7), the axes of A, B, C and of B, C, D have common perpendiculars in $\mathbb{H}^2$.

Example 7.2.3. Suppose that at the level of Möbius transformations, W is a word in the letters A, B with the property that negating the exponents of all the letters ($A \mapsto A^{-1}$, $B \mapsto B^{-1}$) changes W to W^{-1}. Then $\varphi W \varphi = W^{-1}$, in other words φ interchanges the fixed points of W, if W is loxodromic. The axis of W is orthogonal to the axis of φ. If A, B generate a fuchsian group, then the axes of A, B and W in $\mathbb{H}^2$ necessarily intersect at a fixed point of φ. Conclude with [Jørgensen 1978] that on any hyperbolic Riemann surface, a point x which is at the intersection of two closed geodesics, or is at the intersection of a closed geodesic with itself, is at the intersection of infinitely many distinct closed geodesics.

Proof of Proposition 7.2.1. (1) This is verified by a direct matrix computation. One may assume that A is either a diagonal matrix (elliptic or loxodromic), or one with a zero in the lower left entry and trace two (parabolic). The second statement follows from (7.5). See also Lemma 1.5.2.

(2) It follows from (7.1) that

$$(AB - BA) - (BA^{-1} - A^{-1}B) = (A + A^{-1})B - B(A + A^{-1}) = 0.$$

Similarly,

$$(AB - BA) - (B^{-1}A - AB^{-1}) = A(B + B^{-1}) - (B + B^{-1})A = 0.$$

Consequently,

$$\varphi = AB - BA = BA^{-1} - A^{-1}B = B^{-1}A - AB^{-1}.$$

Since

$$(AB - BA)A^{-1} = A(BA^{-1} - A^{-1}B), \quad (AB - BA)B^{-1} = B(B^{-1}A - AB^{-1}),$$

Equations (7.6) follow.

In the remainder of the proofs, we have to be careful when switching between matrices and Möbius transformations.

(3) Since its trace is zero, φ is a half-rotation. The relations (7.6) show that φ interchanges the fixed points or fixes the fixed point of A (and B) according to whether there are two or one. In the former case, the rotation axis of φ is orthogonal to the axis of A. In the latter case, the fixed point is an endpoint of the axis of φ. In all cases, there is only one line in $\mathbb{H}^3$ with the properties of the axis of φ. The last statement follows from (7.3).

(4) We have $\det \psi = 4 - \tau_A^2 \neq 0$ since $A \neq \pm I$ is not parabolic. Now ψ has zero trace and $\{\psi, A\} = 0$.

(5) If A and B are themselves half-rotations and hence equal to their inverses, (7.6) implies that $\{\varphi, AB\} = 0$. Hence by (1), the fixed points of AB on $\partial\mathbb{H}^3$ are the same as those of φ. If the axes of A and B are known to intersect at a point $x \in \mathbb{H}^3$, AB necessarily fixes x as well. Thus AB is elliptic with the same rotation axis as φ. If in addition AB is a half rotation, then from (7.1), $AB = -BA$ so $\varphi = 2AB$. Correspondingly, $\{B, AB\} = 2A$ and $\{AB, A\} = 2B$. The three axes of A, B, AB intersect mutually orthogonally at x.

(6) If A is not parabolic, its matrix is conjugate to a diagonal matrix. The matrix for B has the form

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix}, \quad \text{with } -a^2 - bc = 1.$$

The axis of B is orthogonal to that of A if and only if $a = 0$. This is exactly the condition that $\tau_{AB} = 0$. In this case, the axis of AB is also orthogonal to that of A.

If A is parabolic, then $\tau_{AB} = 0$ if and only if the fixed point of A is an endpoint of the axis of B. In this case AB is a half-rotation fixing the fixed point of A.

(7) From (7.1),

$$\tau_{ABC} I - \tau_{BAC} I = (AB - BA)C + C^{-1}(B^{-1}A^{-1} - A^{-1}B^{-1}).$$

Moreover,

$$\tau_{AB} I = AB + B^{-1}A^{-1} = \tau_{BA} = BA + A^{-1}B^{-1},$$

and therefore,

$$B^{-1}A^{-1} - A^{-1}B^{-1} = BA - AB = -\varphi.$$

Consequently,

$$\varphi C - C^{-1}\varphi = (\tau_{ABC} - \tau_{CBA})I.$$

When $\tau_{ABC} = \tau_{CBA}$, the axis of φ, which is already known to be orthogonal to that of A and B, is also orthogonal to the axis of C, since $\varphi C \varphi^{-1} = C^{-1}$. Conversely, the only line orthogonal to the axis of both A and B is the axis of φ. If that is also orthogonal to the axis of C, then $\varphi C \varphi^{-1} = C^{-1}$ and $\tau_{ABC} = \tau_{CBA}$.

Equation (7.7) is satisfied if there is a linear relation between A, B, C.

Finally, if all three of A, B, C are half-rotations,

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1} = -CBA.$$

Therefore,

$$\tau_{ABC} - \tau_{CBA} = 2\tau_{ABC},$$

and $\tau_{ABC} = \tau_{CBA}$ if and only if $\tau_{ABC} = 0$, that is, if and only if ABC is a half-rotation.

□

7.3 Square roots

The normalized matrix B is a *square root* of the normalized matrix A if $B^2 = A$. We will use the notation $B = \sqrt{A}$ or $B = A^{1/2}$ with the understanding that the roots are determined only up to the factor ± 1.

Lemma 7.3.1. *We have*:

$$\sqrt{A} = \pm\frac{A+I}{\sqrt{2+\tau_A}} \quad \text{if } \tau_A \neq -2,$$
$$\sqrt{-A} = \pm\frac{A-I}{\sqrt{2-\tau_A}} \quad \text{if } \tau_A \neq 2.$$

The square roots of $-I$ are the normalized matrices with zero trace.

At the level of Möbius transformations, if $A \neq \mathrm{id}$ is not parabolic it has two square roots $A+I$ and $A-I$. If $\sqrt{A}$ denotes one of them, the other has the form $\boldsymbol{\alpha}\sqrt{A}$ where $\boldsymbol{\alpha}$ is the half-rotation about the axis of A.

If A is parabolic, at the level of Möbius transformations A has one root, namely either $A+I$ or $A-I$ depending on whether τ_A is $+2$ or -2. It is parabolic as well.

Proof. It follows from (7.1) that $(A \pm I)^2 = (\tau_A \pm 2)A$. It is also true that $\det(A \pm I) = 2 \pm \mathrm{tr}_A$. In terms of normalized matrices,

$$\left(\frac{A+I}{\sqrt{2+\tau_A}}\right)^2 = A, \quad \left(\frac{A-I}{\sqrt{2-\tau_A}}\right)^2 = -A,$$

where one or the other formula holds if $\tau_A = \pm 2$. If $\tau = \tau_A \neq \pm 2$, again using (7.1),

$$\frac{A \pm I}{\sqrt{2 \pm \tau}} = \mp\frac{A - A^{-1}}{\sqrt{4-\tau^2}}\frac{A \mp I}{\sqrt{2 \mp \tau}}.$$

Now interpret the matrices as the corresponding Möbius transformations. In view of the equation above and Proposition 7.2.1(4), application of the half-rotation ψ about the axis of A sends one root of A to the other.

Finally suppose $B^2 = -I$. Because $\mathrm{tr}(B^2) = -2$ we may conjugate the matrix B so as to have the form $\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$, with $ad = 1$ and $a^2 + d^2 = -2$. These two equations imply that $\tau_B = 0$. We already know that any normalized matrix B of zero trace has the property that $B^2 = -I$. Conversely, any matrix of trace zero is a half-rotation. □

As an application we display the matrix formula

$$\{A, B\} = \sqrt{2-\tau_K}\sqrt{-K}\,BA, \quad K = ABA^{-1}B^{-1}. \tag{7.9}$$

7.4 Complex distance

If $T : z \mapsto ke^{i\theta}z$, $k > 1$, in normalized form, the term *complex length* $\mathcal{L}$ of the resulting geodesic in $\mathbb{H}^3/\langle T \rangle$ refers to the number

$$\mathcal{L} = \log k + i\theta \quad \mathrm{mod}(2\pi i).$$

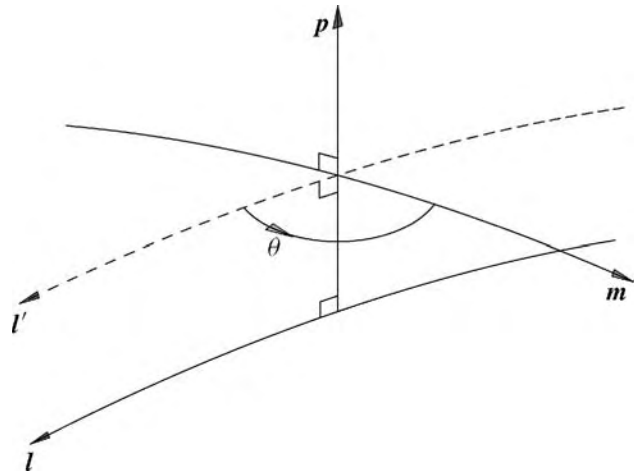


Fig. 7.1. Complex distance. The lines $\boldsymbol{l}$ and $\boldsymbol{l}'$ are coplanar. θ is the angle through which one must rotate $\boldsymbol{l}'$ counterclockwise, when looking along direction of the p.

In invariant form the formula is

$$\tau_T = 2\cosh\frac{\mathcal{L}}{2} \quad \mathrm{mod}(\pm 1).$$

Let $\boldsymbol{l}, \boldsymbol{m}$ be two *oriented* (hyperbolic) lines which do not intersect in $\partial\mathbb{H}^3$. The two lines have a unique common perpendicular line $\boldsymbol{p}$. Orient $\boldsymbol{p}$. The *complex distance* between $\boldsymbol{l}$ and $\boldsymbol{m}$ is the number

$$\chi(\boldsymbol{l}, \boldsymbol{m}) = d(\boldsymbol{l} \cap \boldsymbol{p}, \boldsymbol{m} \cap \boldsymbol{p}) + i\theta \pmod{2\pi i}.$$

Here d denotes the signed hyperbolic distance between $\boldsymbol{l} \cap \boldsymbol{p}$ and $\boldsymbol{m} \cap \boldsymbol{p}$; it is positive if the segment from $\boldsymbol{l} \cap \boldsymbol{p}$ to $\boldsymbol{m} \cap \boldsymbol{p}$ runs in the positive direction with respect to the positive direction along $\boldsymbol{p}$.

Thus in the upper half plane model $d(i, 2i) = \log 2 = -d(2i, i)$ if the vertical is oriented toward ∞. If the lines $\boldsymbol{l}$ and $\boldsymbol{m}$ are disjoint, the sign of $\mathrm{Re}\,\chi(\boldsymbol{l}, \boldsymbol{m})$ is positive or negative depending on the orientation of the common perpendicular $\boldsymbol{p}$. If the lines intersect, $\mathrm{Re}\,\chi(\boldsymbol{l}, \boldsymbol{m}) = 0$ and the orientation of $\boldsymbol{p}$ has no effect.

The angle θ is determined as follows. In the plane σ spanned by $\boldsymbol{l}$ and $\boldsymbol{p}$, let $\boldsymbol{l}' \subset \sigma$ denote the line through $\boldsymbol{m} \cap \boldsymbol{p}$, orthogonal to $\boldsymbol{p}$, and oriented parallel to $\boldsymbol{l}$. As sighted along the ray of $\boldsymbol{p}$ from $\boldsymbol{l} \cap \boldsymbol{p}$, let θ be the angle of *clockwise* rotation necessary to rotate the *positive* ray of $\boldsymbol{l}'$ onto the *positive ray* of $\boldsymbol{m}$.

Changing the order $(\boldsymbol{l}, \boldsymbol{m})$ to $(\boldsymbol{m}, \boldsymbol{l})$ changes the angle from θ to $2\pi - \theta$. It also reverses the sign of the distance, the orientation of $\boldsymbol{p}$ being fixed. Therefore

$$\chi(\boldsymbol{m}, \boldsymbol{l}) = -\chi(\boldsymbol{l}, \boldsymbol{m}) \pmod{2\pi i}$$

If $\boldsymbol{l}^-, \boldsymbol{m}^-$ denote $\boldsymbol{l}, \boldsymbol{m}$ with the opposite orientations,

$$\chi(\boldsymbol{l}^-, \boldsymbol{m}) = \chi(\boldsymbol{l}, \boldsymbol{m}^-) = \chi(\boldsymbol{l}, \boldsymbol{m}) + \pi i \pmod{2\pi i}.$$

The two lines $\boldsymbol{l}, \boldsymbol{m}$ lie in a plane if and only if θ is 0 or π, mod 2π.

Complex distance can also be expressed in terms of a cross ratio. Let $\boldsymbol{l}$ be oriented from endpoint r to endpoint s and $\boldsymbol{m}$ from u to v. Then

$$\cosh^2\left(\frac{\chi(\boldsymbol{l}, \boldsymbol{m})}{2}\right) = (r, u, v, s).$$

It suffices to confirm this formula for two lines in the upper half-space model that are orthogonal to the vertical axis.

We see that $\exp\chi(\boldsymbol{l}, \boldsymbol{m}) = ke^{i\theta}$ is a continuous function of the triple of oriented lines $(\boldsymbol{l}, \boldsymbol{p}, \boldsymbol{m})$, where $\boldsymbol{l}$ and $\boldsymbol{m}$ do not intersect and $\boldsymbol{p}$ is orthogonal to $\boldsymbol{l}$ and $\boldsymbol{m}$.

7.5 Complex distance and line geometry

From Section 7.2 we know that the common perpendicular $\boldsymbol{p}$ to lines $\boldsymbol{l}, \boldsymbol{m}$ is the axis of the half-rotation corresponding to the Lie product $\varphi = \{L, M\}$ of the half-rotations L, M about the lines. The line $\boldsymbol{p}$ is also the axis of ML. In fact, we see that the point $\boldsymbol{m} \cap \boldsymbol{p}$ is the midpoint of the segment of $\boldsymbol{p}$ between $\boldsymbol{l} \cap \boldsymbol{p}$ and $ML(\boldsymbol{l}) \cap \boldsymbol{p}$. Therefore

$$\chi(\boldsymbol{l}, \boldsymbol{m}) = \frac{1}{2}\big(d(\boldsymbol{l} \cap \boldsymbol{p}, ML(\boldsymbol{l} \cap \boldsymbol{p})) + 2i\theta\big) = \frac{1}{2}\chi(\boldsymbol{l}, ML(\boldsymbol{l})) \pmod{2\pi i},$$

where d is measured with respect to the positive direction along $\boldsymbol{p}$.

Lemma 7.5.1. *The normalized matrices*

$$M_1 = \frac{ML + I}{\sqrt{\tau_{ML} + 2}}, \quad M_2 = \frac{ML - I}{\sqrt{\tau_{ML} - 2}}$$

correspond to the two transformations with axis $\boldsymbol{p}$, related by a half-rotation, that send $\boldsymbol{l}$ onto $\boldsymbol{m}$.

Proof. By Proposition 7.2.1(1), as Möbius transformations, M_1 and M_2 have the same pair of fixed points and therefore the same axis. Lemma 7.3.1 shows that $M_1^2 = M_2^2 = ML$. The three transformations ML, M_1, M_2 thus share the line $\boldsymbol{p}$ as axis. The orientation of $\boldsymbol{p}$ toward the attracting fixed point of ML agrees with its orientation from $\boldsymbol{l} \cap \boldsymbol{p}$ to $\boldsymbol{m} \cap \boldsymbol{p}$. Both M_1 and M_2 send $\boldsymbol{l}$ to $\boldsymbol{m}$ and $\boldsymbol{l} \cap \boldsymbol{p}$ to $\boldsymbol{m} \cap \boldsymbol{p}$, but they give $\boldsymbol{m}$ opposite orientations. □

We also record the following fact.

Lemma 7.5.2. *Given the axis $\boldsymbol{p}$ of a loxodromic transformation A and any line $\boldsymbol{m}$ orthogonal to $\boldsymbol{p}$, there is a uniquely determined line $\boldsymbol{l}$ also orthogonal to $\boldsymbol{p}$ such that $A = ML$.*

Proof. Orient $\boldsymbol{p}$ toward the attracting fixed point of A and orient $\boldsymbol{m}$ arbitrarily. Set $y = \boldsymbol{m} \cap \boldsymbol{p}$ and find $x \in \boldsymbol{p}$ such that the distance along $\boldsymbol{p}$ from x to y equals the distance from y to Ax. Find $\boldsymbol{l}$ orthogonal to $\boldsymbol{p}$ at x and orient it so that $\chi(\boldsymbol{l}, \boldsymbol{m}) = \frac{1}{2}\chi(\boldsymbol{l}, A\boldsymbol{l})$.

Note that A has automatically has the symmetries $A^{-1} = MAM$ and $A^{-1} = LAL$. □

Consider next the relation of the complex distance $\chi(\boldsymbol{l}, \boldsymbol{m})$ to the trace of a loxodromic $A = ML = \boldsymbol{ml}$. Let λ, λ^{-1} denote the eigenvalues of A with $|\lambda| > |\lambda^{-1}|$ and $\operatorname{Re}\lambda \geq 0$. The Möbius transformation corresponding to A is conjugate to $z \mapsto \lambda^2 z$ and the complex distance along the axis of A, oriented towards the attracting fixed point, from a point x to MLx is $2\log|\lambda| + 2\theta i$, where $\theta = \arg\lambda$, $-\pi/2 < \theta \leq \pi/2$. That is,

$$\begin{aligned} \theta &= \arg\lambda, \quad -\pi/2 < \theta \leq \pi/2, \\ \chi(\boldsymbol{l}, \boldsymbol{m}) &= \log|\lambda| + \theta i, \quad |\lambda| \geq 1, \\ \operatorname{tr}(A) &= 2\cosh\chi(\boldsymbol{l}, \boldsymbol{m}), \quad \operatorname{Re}\operatorname{tr}(A) \geq 0. \end{aligned} \tag{7.10}$$

The transformation A as described by (7.10) is independent of the orientations of $\boldsymbol{l}$ and $\boldsymbol{m}$.

Conversely, given lines $\boldsymbol{l}, \boldsymbol{m}$ and the orthogonal $\boldsymbol{p}$ from $\boldsymbol{l}$ to $\boldsymbol{m}$, define λ by $\log\lambda = \chi(\boldsymbol{l}, \boldsymbol{m}) \pmod{\pi i}$, which is independent of the orientations of $\boldsymbol{l}, \boldsymbol{m}$. Then

$$\tau_{ML} = (\lambda + \lambda^{-1}) = \pm 2\cosh\chi(\boldsymbol{l}, \boldsymbol{m}). \tag{7.11}$$

As a consequence we obtain the two formulas which are independent of the order and orientations of $\boldsymbol{l}, \boldsymbol{m}$ and the orientation of the axis of ML:

$$\tau_{ML}^2 = 4\cosh^2\chi(\boldsymbol{l}, \boldsymbol{m}), \tag{7.12}$$

$$\det(ML - LM) = -4\sinh^2\chi(\boldsymbol{l}, \boldsymbol{m}). \tag{7.13}$$

7.6 Exercises and explorations

7-1. When does $A^2 = 0$? Is $(AB)^{-1/2} = (B^{-1}A^{-1})^{1/2}$?

7-2. Suppose C and C' are orthogonal circles. Write down the equation of the half-rotation J that exchanges C and C' (for example you may choose the center of C to be 0 and the points of intersection to be $\pm ai$, $a > 0$). Suppose C^* is another circle orthogonal to C such that it and its interior is disjoint from C'. Prove that the radius of $J(C^*)$ is less than that of C^*.

7-3. *Ideal tetrahedra.* We will apply Equation 7.3 to the ideal tetrahedron with vertices at ∞ and 0, P, $Q \in \mathbb{C}$ (see Exercise 1-16). Let $M_{0,\infty}$, M_{PQ} denote half rotations about the edges $[0, \infty]$, $[P, Q]$, respectively. Show that

$$M_{0,\infty}M_{PQ} = \frac{1}{(P-Q)}\begin{pmatrix} -(P+Q) & 2PQ \\ 2 & -(P+Q) \end{pmatrix},$$

and denote its trace by τ. Writing $\tau = \lambda + \lambda^{-1}$, where λ denotes the larger eigenvalue, show that

$$\lambda = -\frac{\sqrt{P} \pm \sqrt{Q}}{\sqrt{P} \mp \sqrt{Q}}.$$

The axis of $M_{0,\infty}M_{PQ}$ is orthogonal to the two lines $[0, \infty]$, $[P, Q]$. Let λ denote the larger eigenvalue ($|\lambda| > 1$). Then

$$\log \lambda = \pm \log \frac{-\sqrt{P} - \sqrt{Q}}{\sqrt{P} - \sqrt{Q}} \pmod{\pi i}$$

is the complex distance between the lines $[0, \infty]$ and $[P, Q]$; compare 7.11. In terms of the cross ratio of the endpoints the distance is $\log(-\sqrt{P}, \sqrt{P}, \sqrt{Q}, \infty) \pmod{\pi i}$.

Find the formulas for the distances between the other two pairs of opposite edges of the tetrahedron.

Check your formulas (or derive them in the first place using this case) by applying them to the case both lines are in the upper half plane model of $\mathbb{H}^2$ and there, $P = 1$ and $Q = z^2$. Confirm that the eigenvalue $\lambda = \frac{z+1}{z-1}$; the complex distance between the lines is $\log \lambda$ (what is the ambiguity in these formulas?).

7-4. [Jørgensen 2000] For a loxodromic transformation A, (7.1) can be written $A + A^{-1} = (\lambda + \lambda^{-1})I$, where λ denotes the larger eigenvalue. Show that

$$A^k = -f_{k-1}(\lambda)I + f_k(\lambda)A,$$

where the coefficients are the polynomials

$$f_k(\lambda) = \frac{\lambda^k - \lambda^{-k}}{\lambda - \lambda^{-1}}.$$

Next show that

$$PA = \lim_{k \to +\infty} \left(\frac{A}{\lambda}\right)^k = \frac{A - \lambda^{-1} I}{\lambda - \lambda^{-1}}$$

where the right side is a singular matrix. Also,

$$P(BAB^{-1}) = B(PA)B^{-1}, \qquad (PA)^2 = PN, \qquad PA + PA^{-1} = I,$$
$$(PA)(PA^{-1}) = 0, \qquad A = \lambda PA + \lambda^{-1} PA^{-1}.$$

In particular P has the properties of a projection.

At the level of Möbius transformations, if p_+, p_- denote the attracting and repelling fixed points of A, $\lim A^k(z)/\lambda^k = p_+$ for all $z \neq p_-$. In fact, by first confirming the formula when $p_+ = \infty$, $p_- = 0$, and then using conjugation, PA is the singular matrix

$$PA = \frac{1}{p_+ - p_-} \begin{pmatrix} p_+ & -p_+ p_- \\ 1 & -p_- \end{pmatrix}.$$

7-5. Prove that the transformation corresponding to the matrix B has the same axis as that corresponding to A if and only if $B = bI + aA$ for some scalars a, b. Conclude that all powers $B = A^k$ of A can be so represented.

Suppose the matrices A, B represent half-rotations. Then a matrix C also representing a half-rotation can be represented as $C = aA + bB$ for some scalars a, b if and only if the axis corresponding to C is perpendicular to the common perpendicular $\{A, B\}$ of the axes corresponding to A and B.

7-6. Suppose $G = \langle X, Y\rangle$ is a 2-generator group. Suppose some word $W(X, Y)$ in the letters $X^{\pm 1}, Y^{\pm 1}$ satisfies $W(X, Y) = \mathrm{id}$. Then also $W(X^{-1}, Y^{-1}) = \mathrm{id}$. If X and Y are conjugate, that is if they have the same trace, then $W(Y, X) = \mathrm{id}$.

Hint: For any Möbius transformation φ, $W(\varphi X\varphi^{-1}, \varphi Y\varphi^{-1}) = \varphi W(X, Y)\varphi^{-1}$. Try $\varphi = \{X, Y\}$. If now X and Y are conjugate and both loxodromic or elliptic, find the midpoint O of the segment of the axis of φ from the point that it crosses the axis of X to the point it crosses the axis of Y. Draw the line ℓ through O and orthogonal to the axis of φ so that the half-rotation φ_1 about ℓ interchanges the axes of X and Y and, if they are loxodromic, sends the attracting fixed point of X to that of Y. Then $\varphi_1 X\varphi_1^{-1} = Y$ and $\varphi_1 Y\varphi_1^{-1} = X$. If both X and Y are parabolic, the axis of φ runs between their fixed point. Find the point O on it such that the half-rotation φ_1 about a line ℓ through O and orthogonal to the axis of φ interchanges the fixed points of X and Y. (The point O satisfies $d(O, XO) = d(O, YO)$.)

Show that every nonelementary two-generator group $G = \langle A, B\rangle$ has an involution $A \mapsto A^{-1}$, $B \mapsto B^{-1}$. This is determined by the common orthogonal to their axes. If $\mathcal{M}(G)$ is a handlebody, the quotient under the involution is the complement of a 3-bridge knot, the singular set being the knot.

In fact, applying tameness and considering the compact core, show that for every (nonelementary) 2-generator group G, either $\mathcal{M}(G)$ has finite volume, or G is a free group and $\mathrm{Int}(\mathcal{M}(G))$ is homeomorphic to the interior of a handlebody.

7-7. *Quaternions again.* In addition to the identity matrix I, introduce the three normalized half-rotation matrices

$$E = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad J = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}, \quad K = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

as in Exercise 1-28. These matrices satisfy

$$EJ = K, \quad JK = E, \quad KE = J.$$

Furthermore their rotation axes are mutually orthogonal at their common point of intersection in $\mathbb{H}^3$.

The four matrices I, E, J, K form a basis of the 4-dimensional complex vector space of complex 2×2 matrices. For any such matrix X write $X = x_1 I + x_2 E + x_3 J + x_4 K$ and correspondingly for matrices Y and Z. Here is a list of properties with respect to this linear structure:

(i) $\tau_X = 2x_1$ and $\det X = x_1^2 + x_2^2 + x_3^2 + x_4^2$.

(ii) $\varphi = \{X, Y\} = 2\begin{vmatrix} I & J & K \\ x_2 & x_3 & x_4 \\ y_2 & y_3 & y_4 \end{vmatrix}$, $\mathrm{tr}(\varphi Z) = -4\begin{vmatrix} z_2 & z_3 & z_4 \\ x_2 & x_3 & x_4 \\ y_2 & y_3 & y_4 \end{vmatrix}$, and $\mathrm{tr}(XY) = 2(x_1y_1 - x_2y_2 - x_3y_3 - x_4y_4)$.

Assume for the following that X, Y, Z are neither degenerate nor a multiple of I.

(iii) The axes of X and Y are the same if and only if X, Y, I are linearly dependent or, if $\tau_X = \tau_Y = 0$, if and only if X and Y are dependent. If X is parabolic then Y is too and with the same fixed point if and only if X, Y, I are dependent.

(iv) Z and the half-rotation $Z_0 = z_2 E + z_3 J + z_4 K$ have the same axis. Z is parabolic if and only if Z_0 is degenerate. If X and Y are half-rotations, then $\{X, Y\}$ and XY have the same axis.

(v) The axes of X, Y, Z have a common perpendicular if and only if X, Y, Z, I are linearly dependent, or, if $\tau_X = \tau_Y = \tau_Z = 0$, if and only if X, Y, Z are linearly dependent.

Also work out the formulas for XY and YX in terms of I, E, J, K. What is the condition for $XY = YX$?

7-8. Prove that lines $\boldsymbol{l}, \boldsymbol{m} \in \mathbb{H}^3$ lie in the same plane if and only if $\cosh \chi(\boldsymbol{l}, \boldsymbol{m})$ is a real number.

7-9. Given an ideal tetrahedron (Exercise 1-16) with edges $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}$, construct the three common perpendiculars between opposite edges. A half-rotation about any one of them maps the tetrahedron onto itself. Consequently the three perpendiculars must meet at a point; the three half-rotations are the non-zero elements of the tetrahedral symmetry group $\mathbb{Z}_2 \otimes \mathbb{Z}_2$.

7-10. *Tubular neighborhood of a systole.* [Gabai et al. 2003] In a hyperbolic manifold suppose there is a shortest geodesic γ of length $l > 0$. Show that γ is necessarily a simple geodesic. Prove that there is an embedded tubular neighborhood (see Exercise 1-4) about γ of radius at least $l/4$. *Hint:* suppose the assertion is false. Then the tubular neighborhood of radius $l/4$ intersects itself at a point p. There at least two perpendicular segments from p to γ. Construct a shorter curve by taking a segment of γ together with the two perpendiculars. The tubular neighborhood about γ is the projection of the tubular neighborhood about any lift $\ell \in \mathbb{H}^3$ of γ.

7-11. *A parametrization of* 2*-generator groups.* Here is a way to parametrize two-generator groups $\langle A, B\rangle$ where A, B are loxodromic without a common fixed point. Let ℓ denote the axis of A. Assume first the two lines $\ell, B^{-1}(\ell)$ do not intersect. Find a line $\boldsymbol{\beta}$, orthogonal to the common perpendicular of $\ell, B^{-1}(\ell)$, with the property that the half-rotation $\boldsymbol{\beta}$ exchanges ℓ and $B^{-1}(\ell)$ and sends the attracting fixed point $p \in \ell$ of A to $B^{-1}(p)$. The transformation $T = B\boldsymbol{\beta}$ maps ℓ onto itself; ℓ is the axis of T and p is its attracting fixed point.

The aim is to parametrize groups $\langle A, B\rangle$ to be uniquely determined up to conjugation by the three complex distances with positive real parts: $\chi = \chi(B^{-1}(\ell), \ell)$, $\log \lambda_A$, and $\log \lambda_T$, where λ_A, λ_T are the larger eigenvalues of A, T, upon arranging matters so that $\operatorname{Re} \chi > 0$.

Show that the parametrization works even if the lines $\ell, B^{-1}(\ell)$ intersect in $\mathbb{H}^3$ and/or B is elliptic.

This parametrization was used in [Gabai et al. 2003] to computationally explore hyperbolic manifolds with short geodesics. If $\langle A, B\rangle$ is discrete, and A, B arise from simple geodesics on the quotient, one can find elements $Y \in \langle A, B\rangle$ such that for the pair A, Y, $\operatorname{Re}\chi$ is as small as possible (but not 0). This means that in the quotient manifold, the geodesic γ resulting from projecting ℓ is contained in an embedded tubular neighborhood of radius $\operatorname{Re}\chi$ and no larger radius. The authors explore the question: Is $\operatorname{Re}\chi \geq (\ln 3)/2$ for all discrete groups $\langle A, B\rangle$? They reduce their problem to a study of short geodesics in discrete groups determined by parameters that lie in a certain box in six-dimensional euclidean space. This required a massive computation of about three CPU years.

7-12. Consider half-infinite polygonal arcs L in $\mathbb{H}^3$ made up of closed geodesic segments $\ell_0, \ell_1, \ldots$ with the property that each ℓ_i meets ℓ_{i+1} at 90°. Let $\boldsymbol{\alpha}_i$ denote the complex distance between ℓ_{i-1} and ℓ_{i+1}, for $i = 1, 2, \ldots$.

Given a sequence of numbers $\{\boldsymbol{\alpha}_i\} \subset \mathbb{C}$ and a base point $O \in \mathbb{H}^3$, show that such a polygonal line L from O can be uniquely constructed. Investigate under what circumstances depending on L that the sequence $\{\boldsymbol{\alpha}_i\}$ converges. When is the sequence dense in $\mathbb{H}^3$?

7-13. *The McShane identity*. Greg McShane [1998] made the following remarkable discovery. On any once punctured torus $\mathbb{T}$,

$$\sum_{\sigma \in \mathcal{S}} \frac{1}{1+e^{\ell(\sigma)}} = \frac{1}{2},$$

where $\mathcal{S}$ is the collection of simple closed geodesics on $\mathbb{T}$ and $\ell(\sigma)$ is the length of σ.

There is an interesting generalization in [McShane 1998; Bowditch 1996]. Take a quasifuchsian representation $\rho : \pi_1(\mathbb{T}) \to \Gamma$ onto a once-punctured torus group. For each $\sigma \in \mathcal{S}$, let $\ell(\rho(\sigma))$ now denote the complex length of $\rho(\sigma)$ modulo $2\pi i\mathbb{Z}$. Then

$$\sum_{\sigma \in \mathcal{S}} \frac{1}{1+e^{\ell(\rho(\sigma))}} = \frac{1}{2},$$

where the sum converges absolutely.

7-14. An orientation reversing isometry J of $\mathbb{H}^3$ with $J^2 = \mathrm{id}$ uniquely determines a plane P_J such that J is the reflection in P_J, possibly followed by the half-rotation in a line in P_J. The latter case does not arise if it is known that J fixes three distinct points in the circle ∂P_J. In other words, J is conjugate on $\partial\mathbb{H}^3$ to $z \mapsto \bar{z}$ or to $z \mapsto -\bar{z}$; see Exercise 1-39.

As an application, suppose J is an orientation reversing isometry of $\mathcal{M}(G)$ with $J^2 = \mathrm{id}$. Assume that $S \subset \mathcal{M}(G)$ is a properly embedded, compact, orientable, incompressible surface, possibly with boundary $\partial S \subset \partial\mathcal{M}(G)$, that is neither a disk nor a cylinder. Assume that there is an arc τ from $O \in S$ to $J(O)$ such that for all loops $\gamma \in \pi_1(S; O)$, $\tau^{-1}J(\gamma)\tau$ is homotopic to γ. Show that there is a totally geodesic surface $S^* \subset \mathcal{M}(G)$ that is pointwise fixed by J. Moreover S^* is homotopic to S.

7-15. *Symmetry lines*. There are two notions of symmetry lines that have been important in studying the combinatorics of fundamental polyhedra: the first for studying cyclic groups [Jørgensen 1973], the second for studying once punctured torus quasi-fuchsian groups [Jørgensen 2003].

Symmetry lines I. Fix a point $\mathcal{O} \in \mathbb{H}^3$ which we will call the basepoint. Given a Möbius transformation g which does not fix $\mathcal{O}$, let $\boldsymbol{\beta}$ denote the line through $\mathcal{O}$ which is orthogonal to the axis of g, if g not parabolic. If g is parabolic, let β be the line through $\mathcal{O}$ to its fixed point. Next construct the plane e_g which is the perpendicular bisector of the line segment $[\mathcal{O},\ g^{-1}(\mathcal{O})]$. In particular, if $\mathcal{O}$ is the origin in the ball model or ∞ in the upper half-space model, $e_g = \mathcal{I}(g)$, the isometric plane.

If instead $\mathcal{O} \in \partial\mathbb{H}^3$, and g does not fix $\mathcal{O}$ construct e_g as follows. There is a unique horosphere $\mathcal{H}_{\mathcal{O}}$ at $\mathcal{O}$ such that the horosphere $g^{-1}\mathcal{H}_{\mathcal{O}}$ at $g^{-1}\mathcal{O}$ is externally tangent to $\mathcal{H}_{\mathcal{O}}$. The line between $\mathcal{O}$ and $g^{-1}\mathcal{O}$ necessarily passes through the point of tangency. (See Lemma 1.5.4.) Take e_g to be the plane through the point of tangency orthogonal to the line between $\mathcal{O}$ and $g^{-1}(\mathcal{O})$.

Recall from Lemma 7.3.1 that if $\mathrm{tr}^2(g) \neq 4$, the square roots of $g^{\pm 1}$ are given as normalized matrices by

$$g^{1/2} = \frac{g \pm I}{\sqrt{2 \pm \tau_g}}, \qquad g^{-1/2} = \frac{g^{-1} \pm I}{\sqrt{2 \pm \tau_g}}.$$

If $\mathrm{tr}^2(g) = 4$ ($g \neq \mathrm{id}$), the roots of $g^{\pm 1}$ are given by the two expressions above that have nonvanishing denominators.

Lemma 7.6.1. *Set $\boldsymbol{\alpha} = g^{-1/2}(\boldsymbol{\beta})$. Then $g = \boldsymbol{\beta}\boldsymbol{\alpha}$ and the line $\boldsymbol{\alpha} \subset e_g$. Furthermore, $e_{\boldsymbol{\alpha}} = e_g$. Let $\boldsymbol{\beta}^*$ denote the line through $\mathcal{O}$ orthogonal to $\boldsymbol{\alpha}$. The plane e_g can be alternately characterized as that plane orthogonal to $\boldsymbol{\beta}^*$ at its point of intersection with $\boldsymbol{\alpha}$.*

Proof. We begin by remarking that when g is not parabolic, the line $\boldsymbol{\alpha}$ is independent of which square root of g is selected, for $\boldsymbol{\beta}$ is orthogonal to the axis of g.

Assume first that $\mathcal{O} \in \mathbb{H}^3$. In the ball model of $\mathbb{H}^3$, replace g by a conjugate so that $\mathcal{O}$ becomes the origin. Then e_g is the isometric plane $\{\vec{x} : |g'(\vec{x})| = 1\}$. Also $|\boldsymbol{\beta}'(\vec{x})| = 1$ for all $\vec{x} \in \mathbb{H}^3$ since $\boldsymbol{\beta}$ is now a euclidean rotation about a diameter. It suffices to prove Lemma 7.6.1 under this normalization of $\mathcal{O}$.

Suppose first g is not parabolic. Recall from Lemma 7.5.2 and Lemma 7.5.1 that $\boldsymbol{\alpha}$ can be alternately described as the line orthogonal to the axis of g such that g is the composition of the half-rotations $g = \boldsymbol{\beta}\boldsymbol{\alpha}$; the axis of g is the common perpendicular of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. We claim that the axis of $\boldsymbol{\alpha}$ lies in e_g and is orthogonal to the segment $[\mathcal{O}, g^{-1}(\mathcal{O})]$ at its midpoint.

For any $\vec{x} \in \mathbb{H}^3$, $|g'(\vec{x})| = |\boldsymbol{\alpha}'(\vec{x})|$. At any fixed point $\vec{x}$ of $\boldsymbol{\alpha}$, $|\boldsymbol{\alpha}'(\vec{x})| = 1$. Thus the axis of $\boldsymbol{\alpha}$ lies in the plane e_g. (Another argument is that $\boldsymbol{\alpha}$ maps e_g onto itself, and its action in e_g is conjugate to the action $z \to \bar{z}$ in $\mathbb{C}$ so that $|\boldsymbol{\alpha}'(\vec{x})| = 1$ for all $\vec{x} \in e_g$.)

Now $g^{-1}(\mathcal{O}) = \boldsymbol{\alpha}\boldsymbol{\beta}(\mathcal{O}) = \boldsymbol{\alpha}(\mathcal{O})$. Since $\boldsymbol{\alpha}$ maps the segment $[\mathcal{O}, g^{-1}(\mathcal{O})]$ onto itself switching the endpoints, it fixes the midpoint $\vec{p}$. Therefore $[\mathcal{O}, g^{-1}(\mathcal{O})]$ must be orthogonal to the axis of $\boldsymbol{\alpha}$ at $\vec{p}$. Since $\mathcal{O}$ and $g^{-1/2}(\mathcal{O})$ have the same distance from the axis of g, and since $\boldsymbol{\alpha} = g^{-1/2}(\boldsymbol{\beta})$, necessarily $\vec{p} = g^{-1/2}(\mathcal{O})$. Finally, $e_{\boldsymbol{\alpha}}$ is the perpendicular bisector of the segment $[\mathcal{O}, \boldsymbol{\alpha}(\mathcal{O}) = g^{-1}(\mathcal{O})]$ so that $e_{\boldsymbol{\alpha}} = e_g$.

If instead g is parabolic, the line $\boldsymbol{\alpha}$ also goes through the fixed point of g. It is also true that $g = \boldsymbol{\beta\alpha}$. The rest of the argument is the same. Another way to confirm this is to assume g first is loxodromic but then allow it to converge to a parabolic and follow the geometry.

Upon reviewing the proof we can confirm that there is no essential difference if $\mathcal{O} \in \partial\mathbb{H}^3$. In this case however it is an endpoint of $\boldsymbol{\alpha}$ that is $g^{-1/2}(\mathcal{O})$. The line $\boldsymbol{\alpha}$ cuts the line $[\mathcal{O}, g^{-1}(\mathcal{O})]$ at its intersection with e_g. □

The line $\boldsymbol{\alpha} \subset e_g$ is called the *symmetry line* of the plane e_g.

Both the plane e_g and its symmetry line $\boldsymbol{\alpha} = \boldsymbol{\alpha}_g$ are uniquely determined by g once $\mathcal{O}$ is chosen (and is not a fixed point of g). In the ball model, if $\mathcal{O} = 0$, then e_g is the isometric plane. In the upper half-space model, when $\mathcal{O} = \infty$, e_g is the isometric plane. Therefore using the ball model, if $\mathcal{O} \in \mathbb{H}^3$ then g can be replaced by AgA^{-1} where $A\mathcal{O} = 0$ and Ae_g is the isometric plane for AgA^{-1}. Using the upper half-space model, if $\mathcal{O} \in \partial\mathbb{H}^3$, then g can be replaced by AgA^{-1} where $A\mathcal{O} = \infty$. These observations are important enough to record formally:

Lemma 7.6.2. *Let $\mathcal{O} \in \mathbb{H}^3 \cup \partial\mathbb{H}^3$ be a basepoint and g a Möbius transformation that does not fix $\mathcal{O}$. Let $\mathcal{O}_1 = A(\mathcal{O})$ be a new basepoint, also not fixed by g. Set $g_1 = AgA^{-1}$. Then the symmetry line $\boldsymbol{\alpha}_1$ and plane e_{g_1} for g_1 with respect to the basepoint $\mathcal{O}_1$ are related to those for g with respect to $\mathcal{O}$ as follows: $\boldsymbol{\alpha}_1 = A(\boldsymbol{\alpha})$ and $e_{g_1} = A(e_g)$.*

Symmetry lines II. The basis for another notion of symmetry line is the following fact.

Lemma 7.6.3. *Suppose A and B are loxodromic while $K = AB^{-1}A^{-1}B$ is parabolic with fixed point $\mathcal{O} \in \partial\mathbb{H}^3$. At the level of Möbius transformations the following hold.*

- *The line $\boldsymbol{\alpha} = \{AB, B^{-1}\} = K^{-1/2}A$ lies in the plane $e_A = e_{\boldsymbol{\alpha}}$.*
- *The line $\boldsymbol{\beta} = \{BA^{-1}B, B^{-1}A\} = K^{-1/2}B$ lies in the plane $e_B = e_{\boldsymbol{\beta}}$.*
- *The line $\boldsymbol{\gamma} = \{B^{-1}AB, B\} = K^{-1/2}AB$ lies in the plane $e_{AB} = e_{\boldsymbol{\gamma}}$.*
- *The line $\boldsymbol{\gamma}^\sharp = \{A, B^{-1}\} = K^{-1/2}AB^{-1} = K^{1/2}B^{-1}A$ lies in the plane $e_{AB^{-1}} = e_{B^{-1}A} = e_{\boldsymbol{\gamma}^\sharp}$.*
- *The lines $\alpha^\sharp = \{AB^{-1}, A^{-1}BA^{-1}\} = K^{1/2}A^{-1}$, $\beta^\sharp = \{A, B^{-1}A^{-1}\} = K^{1/2}B^{-1}$ are the symmetry lines of $e_{A^{-1}}$, $e_{B^{-1}}$ respectively.*
- $B = \boldsymbol{\alpha\gamma} = \boldsymbol{\gamma}^\sharp\boldsymbol{\alpha}, \quad K^{-1}A = \boldsymbol{\beta\gamma} = \boldsymbol{\beta}^\sharp\boldsymbol{\gamma}^\sharp, \quad \boldsymbol{\beta\gamma\alpha} = \boldsymbol{\alpha\gamma}^\sharp\boldsymbol{\beta}^\sharp = K^{-1/2}$.
- *The lines $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\gamma}^\sharp, \ldots$ are mutually disjoint in $\mathbb{H}^3 \cup \mathbb{S}^2$.*

Proof. Replace the transformations by conjugates if necessary so that the point $\mathcal{O}$ becomes ∞ in the upper half-space model. Necessarily $\tau_K = -2$, since $\{A, B^{-1}\}$ is nonsingular. At the level of transformations and referring to Lemma 7.3.1, $K^{1/2} = K - I$ and $K^{-1/2} = K^{-1} - I$; they are parabolic transformations fixing ∞. The formulas are verified by using $K-I=(AB^{-1}-B^{-1}A)A^{-1}B=AB^{-1}(A^{-1}B-BA^{-1})$ and correspondingly expressing $K^{-1}-I$. We also use the facts that the any half-rotation is identical to its inverse and that X, $K^{\pm 1/2}X$, and $K^{\pm 1}X$ all have the same isometric circle and plane.

Finally if $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$, or any two distinct symmetry lines, intersected at $\vec{x} \in \mathbb{H}^3$, then B would fix x and could not be loxodromic. If instead the two half rotations $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ had a common fixed point on S^2, then the composition $\boldsymbol{\alpha\gamma}$ would have zero trace, which is impossible. □

In the situation of Lemma 7.6.3, the lines $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ will be called the *symmetry lines* for the planes e_A, e_B, e_{AB},... respectively. Correspondingly, $\boldsymbol{\alpha}, \boldsymbol{\beta}^\sharp, \boldsymbol{\gamma}^\sharp$ are the symmetry lines for $\mathcal{I}(A), \mathcal{I}(B^{-1}), \mathcal{I}(AB^{-1})$.

Lemma 7.6.2 is worth repeating to cover the present case.

Lemma 7.6.4. *Let $\mathcal{O} \in \mathbb{H}^3 \cup \partial\mathbb{H}^3$ be a basepoint not fixed by A, B and set $\mathcal{O}_1 = X(\mathcal{O})$ and $K_1 = XKX^{-1}$. Then Lemma 7.6.3 holds with respect to the basepoint $\mathcal{O}_1$ with A, B and the half rotations replaced by their conjugates XAX^{-1}, XBX^{-1}.*

7-16. Find the analog of Lemma 7.6.3 in the case that A and B are loxodromic but their commutator is elliptic.

7-17. *Extension of once-punctured torus groups* [Wada 2003]. Prove:

(i) Every quasifuchsian once punctured torus group $G = \langle A, B\rangle$ with parabolic commutator $K^2 = [A, B]$ has an index two extension $G^* = \langle P, Q, R\rangle$ generated by three half-rotations P, Q, R satisfying $RQP = K$.

(ii) Every quasifuchsian twice punctured torus group $G = \langle A, B, L, L'\rangle$ with L, L' both parabolic and $[A, B] = L'L$ has an index two extension $G^* = \langle P, Q, R, S\rangle$ generated by four half-rotations P, Q, R, S satisfying $SRQP = L$.

In the first case the quotient orbifold is $(2, 2, 2, \infty)$, while in the second it is $(2, 2, 2, 2, \infty)$. Here we are referring to punctured spheres with the indicated branching.

Hint: For the first case set $Q = AB - BA$ and define $R = AQ$, $P = BQ$. For the second case define Q, R, P by the same formulas, noting that $RQP = K$ where $K^2 = [A, B]$. Then apply Wada's lemma (Exercise 1-34) to find a half-rotation S with $SLS = L'$.

7-18. *Conformal averaging on $\mathbb{S}^1$* [Schwartz 2006]. Let $W = \{w_1, w_2, \ldots, w_n\}$ be $n \geq 4$ distinct, cyclically arranged points on the unit circle $\partial\mathbb{H}^2$. The complementary intervals $\{(w_i, w_{i+1})\}$ will be subdivided as follows. Construct the common orthogonal ℓ to the two lines $[w_i, w_{i+1}]$, $[w_{i-1}, w_{i+2}]$. One of the endpoints w_i' of ℓ lies in the interval (w_i, w_{i+1}). When this is carried out for all intervals we end up with a new cyclically ordered set of distinct points $W' = \{w_1', w_2', \ldots, w_n'\}$ on the circle. Set up the interactive process $\{W^{(k+1)} = \left(W^{(k)}\right)'\}$. Rich Schwartz proved that $\{W^{(2k)}\}$ converges exponentially fast to an ideal regular n-gon as $k \to \infty$. He interprets this

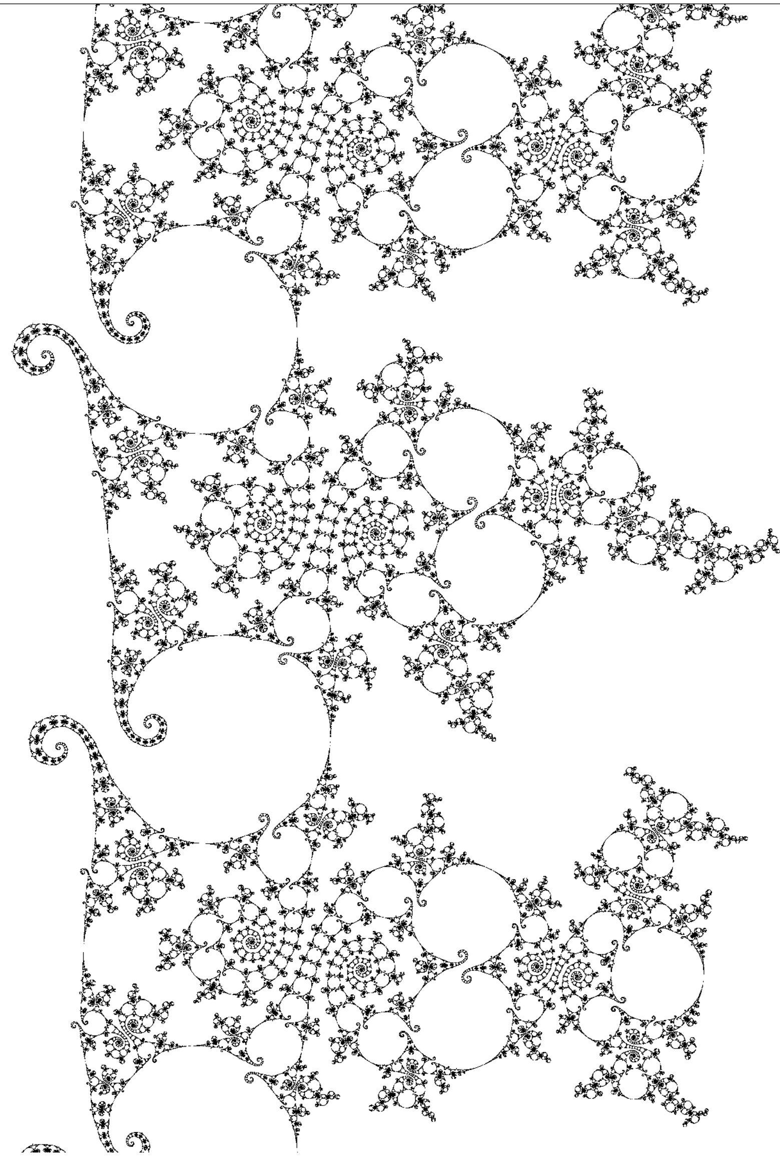


Fig. 7.2. The limit set of a twice punctured torus quasifuchsian group computed using Wada's characterization.

process as "conformal averaging". In the classical situation where w_i' is chosen as the *midpoint* of (w_i, w_{i+1}), the points $W^{(k)}$ become evenly spaced as $k \to \infty$.

7-19. [Brooks and Matelski 1981] Suppose T is a loxodromic with complex translation length δ; that is, if α denotes the axis of T oriented toward its attracting fixed point and ℓ is a line orthogonal to α then $\delta = \chi(\ell, T(\ell))$. Show that

$$\mathrm{tr}^2(T) = 4\cosh^2 \frac{\chi(\ell, T(\ell))}{2},$$

and for any Möbius transformation S with β the axis of $S_1 = STS^{-1}$,

$$\operatorname{tr}(STS^{-1}T^{-1}) - 2 = \big(1 - \cosh\chi(\ell, T(\ell))\big)\big(1 - \cosh\chi(\alpha, \beta)\big).$$

Their paper exploits that the group $\langle S, T\rangle$ is discrete only when $\{\cosh(\chi(\alpha, \beta_i))\}$ is a discrete set. Here β_i the axis of the inductively defined $S_i = S_{i-1}TS_{i-1}{}^{-1}$. More extensive investigations generalizing Jørgensen's inequality are carried out along this line in [Gehring and Martin 1994].

8

Right hexagons and hyperbolic trigonometry

In this chapter we will apply the line geometry developed in Chapter 7 to obtain many of the formulas of hyperbolic trigonometry. Good references are [Beardon 1983] and [Fenchel 1989].

Recall that $\cosh z = (e^z + e^{-z})/2$ and $\sinh z = (e^z - e^{-z})/2$.

8.1 Generic right hexagons

Let $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ be lines in $\mathbb{H}^3$ no two of which have a common point or endpoint, and not all three are orthogonal to the same line. We will use the same notation to represent the half-rotations about these lines. When needed, we will use A, B, C to denote corresponding normalized matrices of zero trace.

The axis $\boldsymbol{\gamma}^*$ of the loxodromic transformation $C_0^* = BA$ is orthogonal to the lines $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, the axis $\boldsymbol{\alpha}^*$ of $A_0^* = CB$ is orthogonal to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, and the axis $\boldsymbol{\beta}^*$ of $B_0^* = AC$ is orthogonal to $\boldsymbol{\gamma}$ and $\boldsymbol{\alpha}$. Note that

$$C_0^* B_0^* A_0^* \;=\; -I, \tag{8.1}$$

so that C_0^*, say, is automatically determined from A_0^* and B_0^*. The half-rotation matrices that correspond to $\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*$ are respectively (see Proposition 7.2.1(4)),

$$A^* = \frac{A_0^* - A_0^{*-1}}{\sqrt{4 - \tau^2(A_0^*)}}, \quad B^* = \frac{B_0^* - B_0^{*-1}}{\sqrt{4 - \tau^2(B_0^*)}}, \quad C^* = \frac{C_0^* - C_0^{*-1}}{\sqrt{4 - \tau^2(C_0^*)}}. \tag{8.2}$$

In terms of Möbius transformations we may write

$$A^* = \{C, B\}, \quad B^* = \{A, C\}, \quad C^* = \{B, A\}. \tag{8.3}$$

The rotation axes of A^*, B^*, C^* are the axes of the loxodromic CB, AC, BA respectively. The intermediate matrices A_0^*, B_0^*, C_0^* are quite useful, irrespective of the awkward notation.

Lemma 8.1.1. (i) *No two of the lines $\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*$ have a common endpoint.*
(ii) *The three lines $\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*$ do not have a common perpendicular.*

Proof. No two of the lines $\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*$ coincide because of our hypothesis that $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ do not have a common perpendicular.

Consider for example $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$. Each one is orthogonal to $\boldsymbol{\gamma}$. If $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$ had a common endpoint then a right-angled hyperbolic triangle with two right angles would be formed. This is impossible since the angle sum must be less than π.

No line ℓ is orthogonal to all three $\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*$. For if ℓ were orthogonal to $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$, say, then they would have two common orthogonals, ℓ and $\boldsymbol{\gamma}$. Because the common orthogonal is a unique line, $\ell = \boldsymbol{\gamma}$. If $\ell = \boldsymbol{\gamma}$ were also orthogonal to $\boldsymbol{\gamma}^*$, then all of $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ would be orthogonal to $\boldsymbol{\gamma}^*$, a contradiction. □

On the other hand it is possible that two of the lines $\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*$ intersect in $\mathbb{H}^3$. If for example $\boldsymbol{\beta}^*$ and $\boldsymbol{\gamma}^*$ intersect at $p \in \mathbb{H}^3$, they span a plane P. The line $\boldsymbol{\alpha}$ is then orthogonal to P at p. This forces the side of the hexagon which lies on $\boldsymbol{\alpha}$ to reduce to the single point p, which is a vertex. We will view such a hexagon as "degenerate". It is possible to avoid such a degeneration by moving one or more of the lines $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ slightly. Degenerate hexagons will be discussed in Section 8.3.

A *generic right hexagon* is one determined by a triple of lines $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ such that

- no pair of lines have a common point or endpoint,
- the three lines do not have a common perpendicular, and
- no two of the dual lines $\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*$ intersect in $\mathbb{H}^3$.

For generic hexagons, the six segments cut off by pairwise intersections in the order

$$(\boldsymbol{\alpha}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}, \boldsymbol{\alpha}^*, \boldsymbol{\gamma}, \boldsymbol{\beta}^*)$$

form a right-angled hexagon $\mathcal{H}ex(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$, which is not planar unless $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ all lie in a plane. Prescribing the cyclic sequence of intersections in the order indicated orients each of the six line segments of the hexagon. The orientation is consistent with the orientation of each line $\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*$ toward the attracting fixed point of each loxodromic transformation $\boldsymbol{\gamma}\boldsymbol{\beta}, \boldsymbol{\alpha}\boldsymbol{\gamma}, \boldsymbol{\beta}\boldsymbol{\alpha}$ respectively. Each side s is opposite its dual side s^*.

We stress that for generic hexagons, the triples of lines $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ and $(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*)$ are interchangeable with each other each other: each triple is dual to the other.

Corollary 8.1.2 (Petersen–Morley Theorem). *The three altitudes of a generic right hexagon in* $\mathbb{H}^3$ *have a common perpendicular.*

Proof. In the notation we have been using, the altitudes are contained in the axes of the three half-rotations $h_\alpha = \{\boldsymbol{\alpha}, \boldsymbol{\gamma}\boldsymbol{\beta}\}$, $h_\beta = \{\boldsymbol{\beta}, \boldsymbol{\alpha}\boldsymbol{\gamma}\}$, and $h_\gamma = \{\boldsymbol{\gamma}, \boldsymbol{\beta}\boldsymbol{\alpha}\}$. These are the common perpendiculars to the pairs of opposite lines $(\boldsymbol{\alpha}, \boldsymbol{\alpha}^*)$, $(\boldsymbol{\beta}, \boldsymbol{\beta}^*)$, $(\boldsymbol{\gamma}, \boldsymbol{\gamma}^*)$, respectively. Now compute the three Lie products; their sum is zero. Apply Proposition 7.2.1(8). □

Let $\mathcal{S}$ denote the space of ordered triples $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ of unoriented lines in $\mathbb{H}^3$ which produce generic right hexagons.

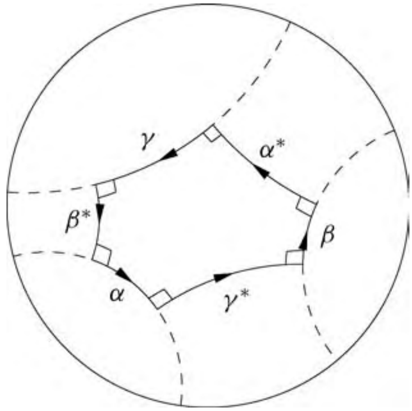


Fig. 8.1. Configuration of sides of the hexagon determined by $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$.

Lemma 8.1.3. *The space $\mathcal{S}$ is connected. Once an initial choice of half-rotation matrices A, B, C corresponding to $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ is made at one point of $\mathcal{S}$, then by continuity a choice is uniquely determined at all other points. In particular one generic right hexagon $\mathcal{H}ex(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ can be moved continuously through generic right hexagons to any other.*

Proof. A line $\ell \in \mathbb{H}^3$ can be moved into a small neighborhood of a point on $\partial\mathbb{H}^3$. Three lines can be moved close to any three distinct points on $\partial\mathbb{H}^3$ without intersecting each other. The three lines can be adjusted so that the hexagon they determine is generic. If in the course of the motion the three lines have a common perpendicular ℓ, than an arbitrarily small change in the position of any one of them (so long as it is not a rotation about ℓ) will destroy this property. Likewise pairwise intersections of the dual lines $\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*$ can be avoided. To be specific, if $\boldsymbol{\gamma}$ is taken as the vertical half-line rising in the upper half space model from from $z = 0$, then the lines $\boldsymbol{\alpha}^*$ with endpoints $u_1, u_2 \in \mathbb{C}$ and $\boldsymbol{\beta}^*$ with endpoints $v_1, v_2 \in \mathbb{C}$ have a common point of intersection with $\boldsymbol{\gamma}$ if and only if $u_2 = -u_1$, $v_2 = -v_1$ while $|u_1| = |v_1|$. To avoid the common intersection all that is needed is to move one endpoint slightly.

Thus the movement to the neighborhood of distinct points on $\partial\mathbb{H}^3$ can be made so that the hexagon remains generic at all intermediate points. □

8.2 The sine and cosine laws for generic right hexagons

In Chapter 7 we introduced the notation of complex distance between *oriented* lines ℓ_1, ℓ_2 as

$$\chi(\ell_1, \ell_2) = \log\rho + i\theta \pmod{2\pi i}, \qquad \rho > 0,$$

where $\log\rho$ is the distance from $\ell_1 \cap \boldsymbol{p}$ to $\ell_2 \cap \boldsymbol{p}$ along the oriented line $\boldsymbol{p}$ orthogonal to ℓ_1 and ℓ_2. We will continue to use the notation from Section 8.1 that $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ are three lines which together with their dual lines $\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*$ form a generic right hexagon. The lines will be oriented as before by choosing the cyclic order of vertices. Since $\sinh(z+\pi i) = -\sinh z$ and $\cosh(z+\pi i) = -\cosh z$ the formulas below are independent of which orientation is chosen. The laws are:

Law of sines
$$\frac{\sinh\chi(\boldsymbol{\beta},\boldsymbol{\gamma})}{\sinh\chi(\boldsymbol{\beta}^*,\boldsymbol{\gamma}^*)} = \frac{\sinh\chi(\boldsymbol{\gamma},\boldsymbol{\alpha})}{\sinh\chi(\boldsymbol{\gamma}^*,\boldsymbol{\alpha}^*)} = \frac{\sinh\chi(\boldsymbol{\alpha},\boldsymbol{\beta})}{\sinh\chi(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*)}. \tag{8.4}$$

Law of cosines
$$\begin{aligned}\cosh\chi(\boldsymbol{\alpha},\boldsymbol{\beta}) &= \cosh\chi(\boldsymbol{\beta},\boldsymbol{\gamma})\cosh\chi(\boldsymbol{\gamma},\boldsymbol{\alpha})\\ &\quad + \sinh\chi(\boldsymbol{\beta},\boldsymbol{\gamma})\sinh\chi(\boldsymbol{\gamma},\boldsymbol{\alpha})\cosh\chi(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*).\end{aligned} \tag{8.5}$$

These formulas apply to any cyclic permutation of the six sides $(\boldsymbol{\alpha},\boldsymbol{\gamma}^*,\boldsymbol{\beta},\boldsymbol{\alpha}^*,\boldsymbol{\gamma},\boldsymbol{\beta}^*)$.

Proof. For a start, we have to confirm the identity

$$4(\operatorname{tr}(XYX^{-1}Y^{-1})-2) = (2\operatorname{tr}(XY)-\operatorname{tr}(X)\operatorname{tr}(Y))^2-(\operatorname{tr}^2(X)-4)(\operatorname{tr}^2(Y)-4). \tag{8.6}$$

This is done by applying Lemma 1.5.6(ii)-(i).

Let $\tilde{C}$ be a normalized matrix corresponding to the loxodromic transformation $\boldsymbol{\beta}^*\boldsymbol{\alpha}^*$. It has axis $\boldsymbol{\gamma}$ and complex translation length $2\chi(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*)$ (7.10). Using (8.3) and (7.13) we find that

$$\pm 4[\sinh\chi(\boldsymbol{\gamma},\boldsymbol{\alpha})\sinh\chi(\boldsymbol{\beta},\boldsymbol{\gamma})]\tilde{C} = (B_0^* - B_0^{*-1})(A_0^* - A_0^{*-1}). \tag{8.7}$$

For later application we will also record the result of using a different formula for the determinants of the right hand side derived in the proof of Proposition 7.2.1(4), namely

$$\pm\left((\operatorname{tr}^2(B_0^*)-4)(\operatorname{tr}^2(A_0^*)-4)\right)^{1/2}\tilde{C} = (B_0^* - B_0^{*-1})(A_0^* - A_0^{*-1}). \tag{8.8}$$

Here A, B, C denote the normalized half-rotations corresponding to $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$.

Expanding the right side of (8.7) and bringing in Lemma 1.5.6(i) we find that the trace of the right side is

$$2\operatorname{tr}(A_0^*B_0^*) - 2\operatorname{tr}(A_0^*B_0^{*-1}) = 4\operatorname{tr}(A_0^*B_0^*) - 2\operatorname{tr}(A_0^*)\operatorname{tr}(B_0^*).$$

Taking the trace of the left side as well,

$$\pm 8\big(\sinh\chi(\boldsymbol{\gamma},\boldsymbol{\alpha})\sinh\chi(\boldsymbol{\beta},\boldsymbol{\gamma})\big)\cosh\chi(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*) = 4\operatorname{tr}(A_0^*B_0^*) - 2\operatorname{tr}(A_0^*)\operatorname{tr}(B_0^*).$$

Squaring both sides gets rid of the $\pm$ ambiguity:

$$16\sinh^2\chi(\boldsymbol{\alpha},\boldsymbol{\gamma})\sinh^2\chi(\boldsymbol{\beta},\boldsymbol{\gamma})\cosh^2\chi(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*) = \big(2\operatorname{tr}(A_0^*B_0^*) - \operatorname{tr}(A_0^*)\operatorname{tr}(B_0^*)\big)^2.$$

Now replace $\cosh^2$ by $1+\sinh^2$. After doing so and after separating the terms on the left, replace $16\sinh^2\chi(\boldsymbol{\alpha},\boldsymbol{\gamma})\sinh^2\chi(\boldsymbol{\beta},\boldsymbol{\gamma})$ by the alternative expression for the determinants of the right hand side of (8.7), as was used in (8.8). In doing so we get our chance to apply (8.6). Applying (8.1) in the process and representing the starred elements in terms of the unstarred ones, the result is:

$$\sinh^2 \chi(\boldsymbol{\alpha}, \boldsymbol{\gamma}) \sinh^2 \chi(\boldsymbol{\beta}, \boldsymbol{\gamma}) \sinh^2 \chi(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) = \frac{\delta^2}{4},$$
$$\delta^2 = \operatorname{tr}(A_0^* B_0^* A_0^{*-1} B_0^{*-1}) - 2 = -\operatorname{tr}(CBA)^2 - 2 = -\operatorname{tr}^2(CBA).$$

Note that δ^2 is invariant under cyclic permutation of A, B, C.

Define δ by choosing the sign of $\pm i \operatorname{tr}(CBA)$ so that

$$\sinh \chi(\boldsymbol{\alpha}, \boldsymbol{\gamma}) \sinh \chi(\boldsymbol{\beta}, \boldsymbol{\gamma}) \sinh \chi(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) = \frac{\delta}{2}. \tag{8.9}$$

We claim the definition of δ for one triple determines it by continuity for all triples of lines $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ without a common point or end point. Start with an initial choice of $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ and matrices A, B, C and let these range over the full space $\mathcal{S}$. The claim is valid simply because at no point can any of the terms in the left side of (8.9) vanish.

The law of sines follows from the identity

$$\frac{\sinh \chi(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)}{\sinh \chi(\boldsymbol{\alpha}, \boldsymbol{\beta})} = \frac{\delta/2}{\sinh \chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) \sinh \chi(\boldsymbol{\beta}, \boldsymbol{\gamma}) \sinh \chi(\boldsymbol{\gamma}, \boldsymbol{\alpha})},$$

since the right side is invariant under cyclic permutation.

We are also ready for the law of cosines. The starting point here is (8.8). Operating on the left side as before, and on the right side bringing in (7.12) we wind up with

$$\pm \sinh \chi(\boldsymbol{\beta}, \boldsymbol{\gamma}) \sinh \chi(\boldsymbol{\gamma}, \boldsymbol{\alpha}) \cosh \chi(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)$$
$$= \cosh \chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) \pm \cosh \chi(\boldsymbol{\beta}, \boldsymbol{\gamma}) \cosh \chi(\boldsymbol{\gamma}, \boldsymbol{\alpha}).$$

We have also used the fact that $B_0^* A_0^* = -AB$, $B_0^* = AC$ and $A_0^* = CB$.

It remains to settle the matter of signs. Once again we do this by continuity. We may assume that $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ determine a planar, regular right hexagon. In this case all the complex distances $\chi(\cdot, \cdot)$ are equal to $d + \pi i$ (why?) where d is the common side length. We now have to look at the possibilities for the equation

$$\pm(1 - \cosh^2 d) = -1 \pm \cosh d.$$

The only situation for which there is a positive solution for $\cosh d$ occurs when the signs are in the order $(+, -)$. The only solution is $\cosh d = 2$ or $d = \log(2 + \sqrt{3})$. This choice gives the law as stated, and it remains true as stated for the whole space by continuity.

As we have seen, the triples of lines $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ and $\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*$ are interchangeable. □

The case of planar right hexagons is assigned as Exercise 8-6.

8.3 Degenerate right hexagons

A line in $\mathbb{H}^3$ is determined by its end points on $\partial\mathbb{H}^3$. A single point $\zeta \in \partial\mathbb{H}^3$ can be regarded as the limit of a sequence of lines, it can be regarded as an *ideal line*. Upon adapting this point of view, it is natural by comparison with (7.3) to represent ζ by the projective equivalence class of the singular, zero-trace, nonzero matrix

$$\begin{pmatrix} \zeta & -\zeta^2 \\ 1 & -\zeta \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \text{ when } \zeta = \infty.$$

The "common perpendicular" γ^* to two distinct lines $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ which have a common endpoint $\zeta \in \partial\mathbb{H}^3$ can be interpreted to be the *ideal line* ζ itself. This ideal line can also be interpreted to be "orthogonal" to the plane spanned by $(\boldsymbol{\alpha}, \boldsymbol{\beta})$. This interpretation agrees with the construction of γ^* as a Lie product. Indeed, if Q_1, Q_2 denote the other endpoints of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ respectively, we have from (7.3) again

$$AB - BA = -\frac{16(Q_1 - Q_2)^2}{(\zeta - Q_1)^2(\zeta - Q_2)^2} \begin{pmatrix} \zeta & -\zeta^2 \\ 1 & -\zeta \end{pmatrix},$$

where A, B denote half-rotation matrices for $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$.

This latter interpretation is the limiting case of two lines which intersect in $\mathbb{H}^3$. Indeed, suppose $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ intersect in $\vec{x} \in \mathbb{H}^3$. The common perpendicular $\boldsymbol{\gamma}^*$ is the line through $\vec{x}$ and perpendicular to the plane spanned by $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. However unlike the case of the common perpendicular between disjoint lines, the ordering $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ no longer determines an orientation of $\boldsymbol{\gamma}^*$. If we choose a sequence of disjoint lines $\boldsymbol{\alpha}_n$, $\boldsymbol{\beta}_n$ which converge to $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, the common perpendicular $\boldsymbol{\gamma}_n{}^*$ of $\boldsymbol{\alpha}_n$ and $\boldsymbol{\beta}_n$ converges to $\boldsymbol{\gamma}^*$. There are two ways to orient $\boldsymbol{\gamma}_n{}^*$ depending on whether $\boldsymbol{\alpha}_n$ is regarded as "over" or "under" $\boldsymbol{\beta}_n$. The two choices induce opposite orientations on $\boldsymbol{\gamma}^*$.

When $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ have a common end point the asymptotic distance between the lines is zero. The complex distance $\chi(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is either 0 or πi depending on the relative orientations of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

When $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ intersect in $\mathbb{H}^3$, they span a plane. We have

$$\chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \pm i\theta \text{ or } \pm i(\pi - \theta),$$

where θ is the acute angle formed by $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ with the sign in each term depending on how $\boldsymbol{\gamma}^*$ is oriented and the choice of term θ or $\pi - \theta$ depending on how $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are oriented to each other. The bottom line is that, in all four cases,

$$\cosh^2 \chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \cos^2\theta, \quad \sinh^2 \chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) = -\sin^2\theta.$$

This also includes the cases of a common endpoint ($\theta = 0, \pi$).

In view of this discussion we will take the expression " common perpendicular" to include the case that one or both "lines" are ideal lines. The common perpendicular between two distinct ideal lines is interpreted as the ordinary line with those end points.

We now return to our construction of right hexagons. Suppose $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ are distinct oriented lines in $\mathbb{H}^3$. We now allow that any two may have a common point or endpoint, but still exclude the possibility that all three have a common perpendicular line (that case is beyond the pale, as no closed figure results). In particular, not all three lines have a common point or endpoint.

Let $\boldsymbol{\gamma}^*$ denote the common perpendicular to $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\alpha}^*$ the common perpendicular to $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\beta}^*$ the common perpendicular to $\boldsymbol{\gamma}$, $\boldsymbol{\alpha}$. Any or all of these perpendiculars may be "ideal lines". At points of intersection in $\mathbb{H}^3$, orient the perpendiculars arbitrarily. The collection of the six oriented lines/ideal lines is called a generalized right hexagon. It is either a generic right hexagon or a it is a *degenerate right hexagon*. Degenerated hexagons have three, four, or five sides; the "degenerate sides" become vertices. If the degenerate side is on $\partial\mathbb{H}^3$, the vertex angle there is zero.

We will treat degenerate hexagons as limiting cases of generic right hexagons, and apply the laws as dictated by continuity.

For example, if $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ form an ideal triangle, $(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*)$ represent the ideal vertices. Conversely if $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ represent three distinct points on $\partial\mathbb{H}^3$, $(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*.\boldsymbol{\gamma}^*)$ are the edges of the associated ideal triangle.

8.4 Formulas for triangles, quadrilaterals, and pentagons

In this section we will present three typical examples of how the right hexagon formulas can be adapted to give formulas for degenerate hexagons, that is, polygons with less than six sides. The trick is to add "ideal" sides of zero length at some vertices so the polygon can then be interpreted as a degenerate case of a right hexagon. Other cases are presented in the exercises.

In each case the edges will be oriented so that the polygonal object lies to the left. In other words, the sides are labeled and oriented so that the vertices appear in the cyclic order $\boldsymbol{\alpha} \cap \boldsymbol{\gamma}^*$, $\boldsymbol{\gamma}^* \cap \boldsymbol{\beta}$, $\boldsymbol{\beta} \cap \boldsymbol{\alpha}^*$, $\boldsymbol{\alpha}^* \cap \boldsymbol{\gamma}$, $\boldsymbol{\gamma} \cap \boldsymbol{\beta}^*$, $\boldsymbol{\beta}^* \cap \boldsymbol{\alpha}$, with appropriate interpretation for degenerate sides.

Right triangles

In (say) the ball model of $\mathbb{H}^3$, move $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ to lie in the equatorial plane forming there a right triangle with the hypotenuse contained in $\boldsymbol{\gamma}$, orientation as shown in Figure 8.4. Then $\boldsymbol{\alpha}^*$, $\boldsymbol{\beta}^*$, $\boldsymbol{\gamma}^*$ are orthogonal to the plane, Orient them so they are pointing toward the lower hemisphere. Denote the side lengths by a, b, c respectively where c is the hypotenuse. Let α, β denote the vertex angles opposite the side lengths a, b respectively.

Then $\chi(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) = c$, $\chi(\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*) = a$, and $\chi(\boldsymbol{\gamma}^*, \boldsymbol{\alpha}^*) = b$ while $\chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \pi i/2$, $\chi(\boldsymbol{\beta}, \boldsymbol{\gamma}) = i(\pi - \alpha)$, and $\chi(\boldsymbol{\gamma}, \boldsymbol{\beta}) = i(\pi - \beta)$. To make this last computation, imagine say $\boldsymbol{\gamma}$ moved slightly out of the plane but parallel to its original position and orthogonal to $\boldsymbol{\beta}^*$. Then looking along $\boldsymbol{\beta}^*$ in the direction of $\boldsymbol{\gamma}$ to $\boldsymbol{\alpha}$ we see that a rotation of angle $\pi - \beta$ is required to rotate $\boldsymbol{\gamma}$ with its orientation onto $\boldsymbol{\alpha}$ with its orientation. Putting these values in the laws gives

$$\cosh c = \cosh a \cosh b, \qquad \sinh c = \frac{\sinh b}{\sin \beta} = \frac{\sinh a}{\sin \alpha}. \tag{8.10}$$

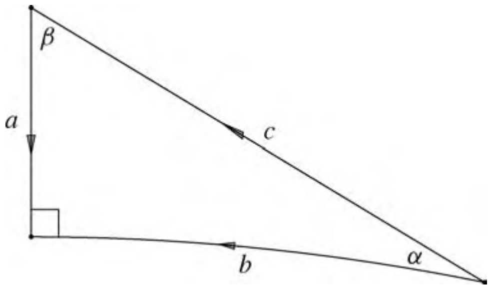


Fig. 8.2. A right (planar) triangle indicating the degenerate sides.

The length of two sides determines the length of the third and also determines the angles. we also find that

$$\tanh b = \sinh a \tan \beta, \quad \operatorname{sech} c = \tan \alpha \tan \beta. \tag{8.11}$$

Some more formulas for right triangles are given in Exercise 8-3. Formulas for the general triangle are presented in Exercise 8-2.

Planar pentagons with four right angles

Move $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ into a plane and orient $\boldsymbol{\gamma}^*$ by thinking of $\boldsymbol{\beta}$ slightly lower than $\boldsymbol{\alpha}$ at their common vertex v so that $\boldsymbol{\gamma}^*$ is oriented so as to point from $\boldsymbol{\alpha}$ to $\boldsymbol{\beta}$. Then $\chi(\boldsymbol{\alpha}, \boldsymbol{\beta})$ approaches θi. The resulting right hexagon has its side on $\boldsymbol{\gamma}^*$ degenerated to the vertex v. We obtain either a convex pentagon or a figure overlapping itself.

Here we will work out the formulas for the convex case; the case of self-intersection is in Exercise 8-5. Place the sides in order $\boldsymbol{\alpha}$, $[\boldsymbol{\gamma}^*]$, $\boldsymbol{\beta}$, $\boldsymbol{\alpha}^*$, $\boldsymbol{\gamma}$, $\boldsymbol{\beta}^*$ where the side on $\boldsymbol{\gamma}^*$ reduces to the vertex v and the interior angle at v is θ. Denote the lengths of the sides contained in $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ by a, b, c, and in $\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*$ by a^*, b^*. Then $\chi(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) = c + \pi i$, $\chi(\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*) = a + \pi i/2$, and $\chi(\boldsymbol{\gamma}^*, \boldsymbol{\alpha}^*) = b + \pi i/2$ while $\chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\pi - \theta)i$. The law of cosines becomes

$$\cosh c \;=\; \sinh a \sinh b - \cosh a \cosh b \cos \theta,$$

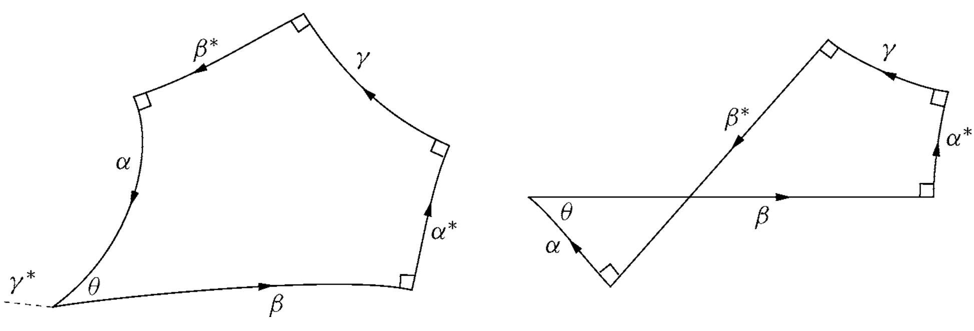


Fig. 8.3. Planar pentagons with four right angles.

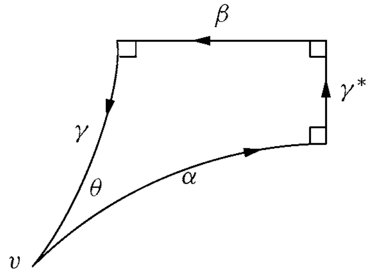


Fig. 8.4. A quadrilateral with three right angles.

or alternatively,

$$\cos\theta = \sinh a^* \sinh b^* \cosh c - \cosh a^* \cosh b^*, \tag{8.12}$$

So c is determined by a, b^* and θ, and θ is determined by a^*, b^*, and c; two convex right pentagons whose corresponding side lengths are identical are isometric.

The law of sines becomes

$$\frac{\cosh a}{\sinh a^*} = \frac{\cosh b}{\sinh b^*} = \frac{\sinh c}{\sin\theta}. \tag{8.13}$$

Specialize to the cases $\theta = 0$ (v is an ideal vertex) and $\theta = \pi i/2$.

Quadrilaterals with three right angles

Here we are not assuming that the quadrilateral is planar. Let v denote the vertex with angle θ. Label the sides in order as $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}^*$, $\boldsymbol{\beta}$, $[\boldsymbol{\alpha}^*]$, $\boldsymbol{\gamma}$, $[\boldsymbol{\beta}^*]$, where the brackets indicate the lines associated with degenerate sides — which correspond to vertices — and $[\boldsymbol{\beta}^*]$ is associated with v.

The lines $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$ are perpendicular to the planes determined by $\boldsymbol{\beta}, \boldsymbol{\gamma}$ and $\boldsymbol{\gamma}, \boldsymbol{\alpha}$ respectively. Orient $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$ to point into these planes, and interpret $\boldsymbol{\beta}$ to lie over $\boldsymbol{\gamma}$ on $\boldsymbol{\alpha}^*$ and $\boldsymbol{\gamma}$ over $\boldsymbol{\alpha}$ on $\boldsymbol{\beta}^*$. Then $\chi(\boldsymbol{\beta}, \boldsymbol{\gamma}) = -\pi i/2$ and $\chi(\boldsymbol{\gamma}, \boldsymbol{\alpha}) = (\theta - \pi)i$. In addition, $\chi(\boldsymbol{\gamma}^*, \boldsymbol{\alpha}^*) = \chi(\boldsymbol{\gamma}^*, \boldsymbol{\gamma}) - \pi i/2$. Using this information the law of sines gives

$$\sinh\chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) \cosh\chi(\boldsymbol{\gamma}^*, \boldsymbol{\gamma}) \;=\; \sin\theta \sinh\chi(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*).$$

The law of cosines tells us that

$$\cosh\chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) \;=\; -\sin\theta \cosh\chi(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*).$$

From the two laws we conclude that

$$\tanh\chi(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) = -\tanh\chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) \cosh\chi(\boldsymbol{\gamma}, \boldsymbol{\gamma}^*), \tag{8.14}$$

and

$$\cos\theta \;=\; \sinh\chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) \sinh\chi(\boldsymbol{\gamma}^*, \boldsymbol{\gamma}). \tag{8.15}$$

In particular if the quadrilateral is planar (and necessarily convex) then

$$\cosh c^* \;=\; \sin\theta \cosh c, \tag{8.16}$$

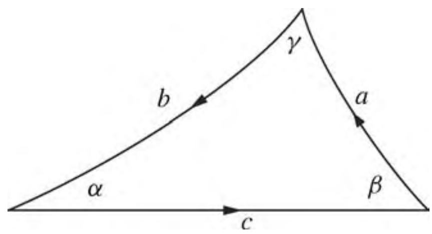


Fig. 8.5. A generic triangle.

where c is the length of a side on $\boldsymbol{\gamma}$ and c^* the length of the opposite side on $\boldsymbol{\gamma}^*$. Equation (8.15) becomes

$$\cos\theta \;=\; \sinh b \sinh c^*, \tag{8.17}$$

where b is the length of the side on $\boldsymbol{\beta}$. This latter equation is used to confirm that the sign is correctly chosen when taking square roots to obtain (8.15).

Two planar quadrilaterals with three right angles with corresponding side lengths identical are isometric. From (8.15) we deduce that $0 \le \theta < \pi/2$; there are no hyperbolic rectangles, as we already know. On the other hand, regular quadrilaterals with $60°$ angles tessellate the plane.

8.5 Exercises and explorations

8-1. Show how to form a right hexagon from six edges of a hyperbolic cube.

8-2. *Law of sines and cosines for triangles.*Consider the general hyperbolic triangle with sides of length a, b, c and opposite angles labeled α, β, γ.

Derive the law of sines and cosines for the triangles, namely:

Law of cosines 8.5.1.

$$\cosh c = \cosh a \cosh b - \sinh a \sinh b \cos\gamma, \tag{8.18}$$

$$\cosh c = \frac{\cos\alpha\cos\beta + \cos\gamma}{\sin\alpha\sin\beta}. \tag{8.19}$$

Law of sines 8.5.2.

$$\frac{\sinh a}{\sin\alpha} = \frac{\sinh b}{\sin\beta} = \frac{\sinh c}{\sin\gamma}. \tag{8.20}$$

Conclude that two triangles with the same angles are isometric.

Show that the quantity $\sinh a \sin\gamma$ is independent of the lengths b, c and the angles α, β that form the triangle. Here is a nice application of this fact:

Suppose $\sigma = \sigma(s) \subset \mathbb{H}^2$ is a geodesic parameterized by hyperbolic arc length and ζ is a point not on σ. At each point $\sigma(s)$ there is a geodesic segment from ζ to $\sigma(s)$;

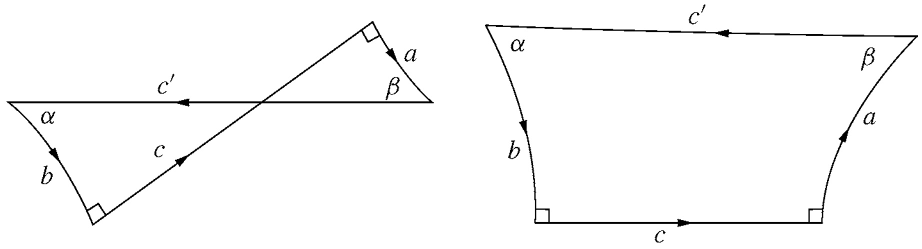


Fig. 8.6. Planar quadrilaterals with two right angles.

denote its length by $x(s)$. Let $\theta(s)$ be the angle of intersection at $\sigma(s)$ measured counterclockwise from σ. Deduce as in [Epstein et al. 2004] the formula

$$\theta'(s) = -\coth x(s)\sin\theta(s).$$

Hint: You will need to know that $x'(s) = \cos\theta(s)$. This is seen using the fact that near $\sigma(s)$ the hyperbolic metric is almost euclidean, therefore $\Delta x \sim \Delta s\cos\theta$.

Finally consider the hyperbolic triangle with vertices x, y, z'. Let z be a point on the side (x, y). Derive the formula

$$\cosh d(z, z')\sinh d(x, y) \;=\; \cosh d(x, z')\sinh d(y, z) + \cosh d(y, z')\sinh d(x, z), \tag{8.21}$$

where as usual $d(\cdot\,,\cdot)$ is the hyperbolic distance.

Prove that if the shortest geodesic γ in a $\mathcal{M}(G)$ has length $\ell > 1.353$, then the tube of radius $\frac{1}{2}\log 3$ about it is embedded (Gabai–Meyerhoff–Thurston). *Hint:* Fix a lift γ^*. If γ_1^* is another of distance d away, then by the law of cosines applied to a right triangle with base along γ_1^* and sides of lengths, $d, \geq \ell, \leq \ell/2$ and $\cosh d \geq (\cosh\ell)/(\cosh\ell/2)$.

8-3. *Right-angled triangles.* Return to the case of aright-angled triangle labeled as in Section 8.4. Verify the additional formulas

$$\cosh c = \cot\alpha\cot\beta,$$
$$\cos\alpha = \tanh b\coth c, \qquad \cos\beta = \tanh a\coth c,$$
$$\sinh a = \cot\beta\tanh\beta, \qquad \sinh b = \cot\alpha\tanh a.$$

Specialize to the case that $b = c = \infty$, $\alpha = 0$.

8-4. *Convex planar quadrilateral with two adjacent right angles.* Let c denote the length of the base which has right angles at its endpoints. Let c' denote the length of the opposite side, and a, b the other two sides. Denote the vertex angle facing side b by β and that facing side a by α; necessarily $\alpha + \beta < \pi$.

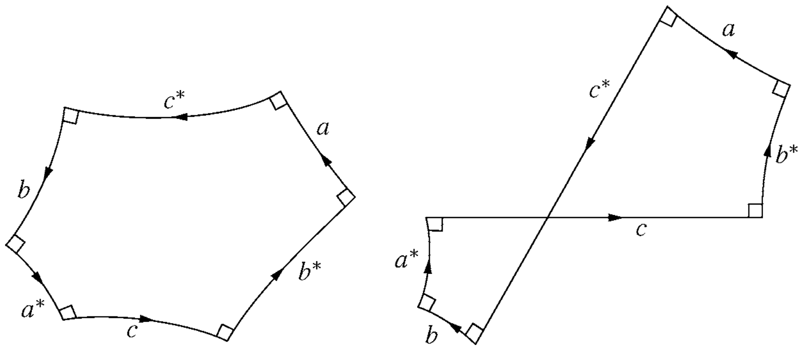


Fig. 8.7. Planar right hexagons.

Verify the formulas

$$\frac{\cosh a}{\sin \alpha} = \frac{\cosh b}{\sin \beta} = \frac{\sinh c'}{\sinh c},$$
$$\cosh c' = -\sinh a \sinh b + \cosh a \cosh b \cosh c$$
$$\cosh c = -\cos \alpha \cos \beta + \sin \alpha \sin \beta \cosh c'.$$

These give the exact formulas for the orthogonal projection in a plane of a line segment onto a geodesic (see Exercise 8-9 for a simpler version).

8-5. Complete the formulas for planar pentagons with four right angles by considering the self-intersecting case when the side along $\boldsymbol{\beta}^*$ crosses the side along b. The law of sines remains the same except for some changes of sign. When applying the law of cosines, for instance $\chi(\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*)$ changes to $a - \pi i/2$.

8-6. *Planar right hexagons.* Work out the formulas in the special case that the right hexagon lies in a plane. Confirm that in the convex case the two laws have the form

$$\frac{\sinh a}{\sinh a^*} = \frac{\sinh b}{\sinh b^*} = \frac{\sinh c}{\sinh c^*}, \tag{8.22}$$

$$\cosh c = -\cosh a \cosh b + \sinh a \sinh b \cosh c^*. \tag{8.23}$$

Derive the corresponding formulas in the self-intersecting case, forming two quadrilaterals each with three right angles.

How many side lengths uniquely determine the hexagon up to isometry, orientation preserving or reversing?

Show that given any $a, b, c > 0$ there exists a unique convex right hexagon with alternating sides of length a, b, c. Equivalently, there exists three mutually disjoint geodesics in $\mathbb{H}^2$ whose respective distances apart are exactly a, b, c.

Specialize to the cases that all sides have the same length, and then that all vertex angles are zero.

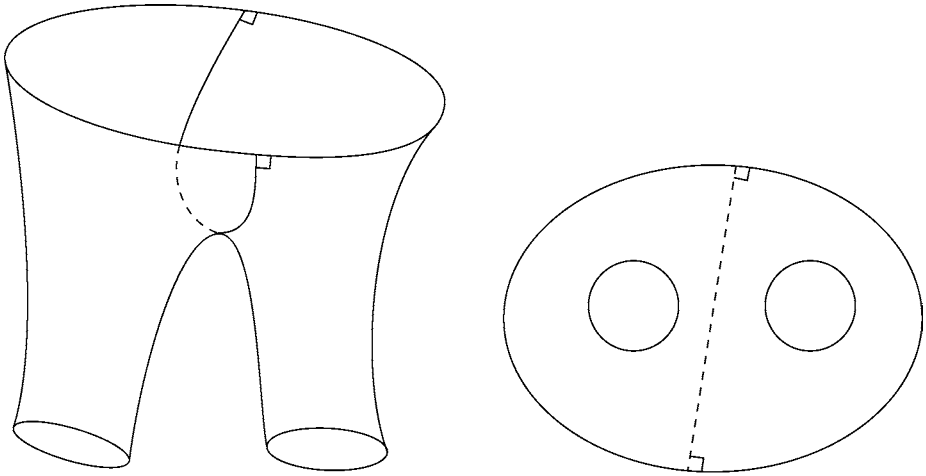


Fig. 8.8. A pair of pants with a seam.

8-7. *Pants.* A pair of pants on a surface is a triply connected planar region bounded by three mutually disjoint closed geodesics. The three common perpendiculars in turn divide the pants into two right hexagons. Show that there is an orientation reversing conformal map (a reflection) that pointwise fixes the common perpendiculars, and interchanges the two hexagons. That is, every pants is preserved by an orientation reversing involution that maps each boundary component onto itself.

In particular the area of a pants is 2π.

Show that the pants is uniquely determined by the lengths of the three boundary geodesics.

In fact suppose the lengths of the three boundary curves c_1, c_2, c_3 of a pants are L_1, L_2, L_3. Denote the length of the seam that is the perpendicular between c_1 and c_2 by ℓ. Show that

$$\cosh \ell = \frac{\cosh L_3 + \cosh L_1 \cosh L_2}{\sinh L_1 \sinh L_2}.$$

See [Fathi et al. 1979, §8] for a description of pants geometry.

Consider the limiting case that one boundary component shrinks to a single point (puncture) and the other two have lengths a, b. There is uniquely determined geodesic whose endpoints are at the puncture and which cuts the pants P into two annular regions P_1, P_2, where P_1 has the a-length side. Each P_i in turn can be cut into two quadrilaterals each with one ideal vertex and three right angles. The two quadrilaterals $Q_{1,i}$, $Q_{2,i}$ of each P_i are symmetric under reflection. One side of $Q_{1,i}$ has length $a/2$, $b/2$ respectively. Denote the lengths of the other two finite sides by a', b', respectively. Show that

$$\sinh(a/2)\sinh a' = 1, \quad \sinh(b/2)\sinh b' = 1.$$

It is a fact (a consequence of the Schwarz Lemma [Ahlfors 1978]) a nested pair of simply connected regions $\Omega_1 \subset \Omega_2$ with corresponding hyperbolic metrics $\rho_1(z)\,|dz|$ and $\rho_2(z)\,|dz|$ have the following property: $\rho_1(z) > \rho_2(z),\ z \in \Omega_1$. From this, deduce the following sharp form of the collar lemma (see [Buser 1992, Chapter 4]).

Collar Lemma 8.5.3. On a hyperbolic surface R suppose $\alpha, \beta \subset R$ are mutually disjoint simple geodesics of lengths a, b. Set

$$a' = \operatorname{arcsinh} \frac{1}{\sinh(a/2)}, \qquad b' = \operatorname{arcsinh} \frac{1}{\sinh(b/2)}.$$

Then the distance-a' annular neighborhood of width $2a'$ about α is disjoint from the distance-b' annular neighborhood of width $2b'$ of β.

For more computational practice show that the length L_α of each boundary component of the collar of distance a' from the geodesic α of length a is

$$L_\alpha = a \cosh a' = \frac{a}{\tanh a/2}.$$

Note that $L_\alpha \to 2$ as $a \to 0$.

Prove [Beardon 1983, Theorem 8.3.1], which asserts that if X, Y generate a nonelementary fuchsian group without elliptic elements then

$$\sinh \frac{d(z, X(z))}{2} \sinh \frac{d(z, Y(z))}{2} \geq 1 \quad \text{for all } z \in \mathbb{H}^2.$$

This inequality is best possible.

8-8. *Polar, cylindrical, and horocyclic coordinates.* Show that for the hyperbolic metric ds in $\mathbb{H}^2$,

$$ds^2 = d\rho^2 + \sinh^2 \rho \, d\theta^2, \tag{8.24}$$
$$ds^2 = \cosh^2 \rho \, dt^2 + d\rho^2, \tag{8.25}$$
$$ds^2 = e^{-2\rho} \, dt^2 + d\rho^2. \tag{8.26}$$

Equation (8.24) is called the polar representation of the hyperbolic metric ds; ρ denotes hyperbolic distance from the origin and θ is the angle from the positive axis to the ray ρ. Equally θ can be interpreted as the angular measure on $\partial\mathbb{H}^2$.

Equation (8.25) is the metric representation in terms of geodesic coordinates; t is arclength along a geodesic α, for example the real diameter in the disk model, and ρ is the distance of a point from α, along an orthogonal line through t. The equation can also be regarded as a two dimensional form of cylindrical coordinates.

Equation (8.26) the metric representation in terms of *horocyclic coordinates*. Here t is arclength along a horocycle, and ρ is signed distance along a geodesic orthogonal at t. We choose the sign of ρ so that positive distance is toward the point on $\partial\mathbb{H}^2$ determined by the horocycle.

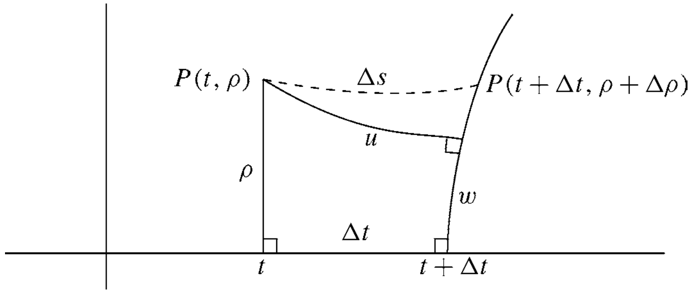


Fig. 8.9. Cylindrical coordinate approximation.

The first formula is derived from expressing the hyperbolic metric in (euclidean) polar coordinates,

$$ds^2 = \frac{4(dx^2 + dy^2)}{(1 - |z|^2)^2} = \frac{4(dr^2 + r^2 d\theta^2)}{(1 - r^2)^2},$$

together with the fact that

$$e^\rho = \frac{1+r}{1-r}.$$

Equation (8.26) is most easily derived in the upper half-plane model with the horocycle $\{y = a > 0\}$. The desired formula results from substituting $x = at$ and $y = ae^\rho$ into $ds^2 = (dx^2 + dy^2)/y^2$.

Equation (8.25) is more complicated to derive. Construct two quadrilaterals Q, Q^* as follows.

The base of Q is on a geodesic; the left endpoint of the base has coordinate t along the geodesic and the right end point is $t + \Delta t$, where Δt is a small deformation. The adjacent vertical sides are at right angles. The left side has length ρ and its top end is labeled $P(t, \rho)$; its coordinates are (t, ρ). The right side has length w. The top side of Q is orthogonal to the right side; denote its length by u. Q has three right angles; the nonright angle is subtended at the vertex $P(t, \rho)$. Now extend Q by taking a small deformation $\Delta\rho$ and extending the right side until it has length $\rho + \Delta\rho$; label the vertex $P(t + \Delta t, \rho + \Delta\rho)$. Let Δs denote the length of the segment from $P(t, \rho)$ to $P(t + \Delta t, \rho + \Delta\rho)$. Thus we now have a larger quadrilateral Q^* whose right side has length $w + v$ for some $v > 0$.

We first work with Q. Insert the diagonal from the upper left to the lower right and denote its length by d. It divides the right angle at the lower right vertex of Q into angles α, $\pi/2 - \alpha$, where we take α to be adjacent to the bottom.

The diagonal forms two right triangles. Therefore

$$\cosh d = \cosh \Delta t \cosh \rho = \cosh u \cosh w,$$

and also

$$\tanh \rho = \sinh \Delta t \tan \alpha; \qquad \tanh u = \sinh w \tan(\pi/2 - \alpha).$$

Now $\tanh u \sim u$ and $\sinh \Delta t \sim \Delta t$, while $\cosh \Delta t \sim 1$ and $\cosh u \sim 1$. Hence $w \sim \rho$, and then

$$u \sim \cosh \rho \Delta t.$$

Now examine the right triangle $Q^* \setminus Q$. The side of length v satisfies $v = \rho + \Delta\rho - w \sim \Delta\rho$. For its hypotenuse, $\Delta s^2 \sim u^2 + v^2$, since the sides are very short. From these calculations Equation (8.25) follows.

The corresponding formulas for $\mathbb{H}^3$ are essentially the same:

$$ds^2 = d\rho^2 + \sinh^2 \rho \, d\theta^2, \tag{8.27}$$

$$ds^2 = \cosh^2 \rho \, dt^2 + d\rho^2 + \sinh^2 \rho \, d\theta^2 \geq \cosh^2 \rho \, dt^2 + d\rho^2, \tag{8.28}$$

$$ds^2 = e^{-2\rho} \, ds^2_{\mathcal{H}} + d\rho^2. \tag{8.29}$$

Equation ((8.29)) is in terms of horocyclic coordinates about a horosphere $\mathcal{H}$, where $ds_{\mathcal{H}}$ denotes distance in $\mathcal{H}$ and ρ is distance along a line orthogonal to $\mathcal{H}$ with positive direction toward the associated point on $\partial\mathbb{H}^3$.

Equation ((8.28)) is in terms of cylindrical coordinates about a geodesic α; t denotes distances along α, $\rho > 0$ denotes distances from α so that (ρ, θ) are polar coordinates in the plane orthogonal to α at t. The volume form in these coordinates is

$$dV = \cosh \rho \sinh \rho \, dt \, d\rho \, d\theta. \tag{8.30}$$

Equation (8.28) invites the following interpretation. Denote the convex core of $\mathcal{M}(G)$ by $\mathcal{C}(G)$. If ρ represents the shortest distance of a point exterior to $\mathcal{C}(G)$ to $\partial\mathcal{C}(G)$, the hyperbolic metric restricted to the exterior of $\mathcal{C}(G)$ satisfies

$$ds^2 \geq \cosh^2 \rho \, ds^2_{\partial\mathcal{C}} + d\rho^2 > \frac{1}{4} e^{2\rho} ds^2_{\partial\mathcal{C}} + d\rho^2.$$

One says that the exterior of $\mathcal{C}(G)$ is *exponentially flaring*.

Boundary length estimates for triangles and cylinders. The hyperbolic area element in polar coordinates is

$$dA = \sinh \rho \, d\rho \, d\theta. \tag{8.31}$$

If Δ is a hyperbolic triangle, v is one of its vertices, and α is the opposite side, Area(Δ) can be expressed in polar coordinates about v as,

$$A = \int\!\!\int_A \sinh \rho(\theta) \, d\rho \, d\theta = \int (\cosh \rho(\theta) - 1) \, d\theta.$$

The area of the triangular region in Δ *outside* the distance-r neighborhood of α is

$$\int (\cosh(\rho(\theta) - r) - 1) \, d\theta < e^{-r} \int \sinh \rho(\theta) \, d\theta < e^{-r} \operatorname{Len}(\alpha).$$

In particular the area of a triangle is less than the length of its shortest side.

Following on in this vein, let $f : S^1 \times [0, 1] \to \mathcal{M}(G)$ be the embedding of a closed cylinder, that is a free homotopy between the simple loops $\gamma_1 = f(\mathbb{S}^1 \times \{0\})$,

$\gamma_2 = f(\mathbb{S}^1 \times \{1\})$. Replace each arc $\{f(\theta\} \times [0, 1])\}$ by a geodesic arc with the same endpoints, obtaining as a consequence a ruled cylinder C. The cylinder C can be approximated by a union of thin quadrilaterals. Each quadrilateral can be divided into two triangles. One of the triangles abuts γ_1, the other abuts γ_2. The total area of the latter triangles is less than the length of γ_2. Following [Thurston 1979, §9.3], prove from the estimate above that

$$\mathrm{Area}(C \setminus N_r(\gamma_2)) \leq e^{-r}\,\mathrm{Len}(\gamma_2) + \mathrm{Len}(\gamma_1), \tag{8.32}$$

where $N_r(\gamma_2)$ is the distance-r annular neighborhood of γ_2. In particular, the area of C is less than the length of ∂C. This finds essential use in the theory of ending laminations for example in Lemma 5.6.7.

8-9. *Orthogonal projection strictly reduces distances.*Establish this often used property: Let $\ell \in \mathbb{H}^3$ be a line and σ be a line segment of finite length which we may assume is disjoint from ℓ. Let $\sigma^* \in \ell$ denote the orthogonal projection of σ to ℓ. Show that the length of σ^* is strictly less than the length of σ (unless σ is itself a segment of ℓ).

That is, if x and y in $\mathbb{H}^3$ lie on the same side of ℓ and have distance $\geq r$ from ℓ, then the orthogonal projection of the two points x, y onto ℓ satisfies

$$d(\pi(x), \pi(y)) < d(x, y). \tag{8.33}$$

Hint: First verify that in a right triangle, the hypotenuse is strictly longer than either leg. Then verify that in a planar quadrilateral with three right angles, the length of a side with one end at the vertex with vertex angle $\theta \neq \pi/2$ is strictly greater than the length of the opposite side. Next show that the same property holds for a nonplanar quadrilateral with three right angles by taking its orthogonal projection to the plane formed by two orthogonal edges. Finally return to the original problem and drop a perpendicular from one end of σ to form a quadrilateral with three right angles and a right triangle.

Deduce from Equation ((8.28)) that it is also true that

$$d(\pi(x), \pi(y)) < 2e^{-r} d(x, y), \tag{8.34}$$

which is better than Equation (8.33) when $r > \log 2$. More precisely, if x, y are equal distance r from ℓ and the segment $[\pi(x), \pi(y)]$ has length L, then

$$2\cosh d(x, y) = 1 + \cosh L + (\cosh L - 1)\cosh 2r.$$

8-10. *Riemann surfaces made out of pentagons.*Show there exists a unique regular right-angled pentagon up to isometry; find a formula for the side length s.

Position it so that one side is the “bottom”. We will refer to the sides directly on its right and its left as the “vertical” sides. Reflect across the bottom giving a 6-sided right polygon with its two vertical sides of length $2s$ and the others of length s. Then reflect this pair across say the right vertical side giving a right angled 8-sided polygon P, the union of four pentagons, four of whose sides have length $2s$. From this you can construct a hyperbolic surface of genus two. From now on, P will be the fundamental unit.

Keep adjoining copies of P to the right side. Each time you attach a unit P, you will attach four pentagons. You will be subtracting an edge and adding five new edges. In short, you will be able to construct a closed surface of genus $g \geq 2$ out of $4(g-1)$ pentagons comprising a 4g sided right polygon. We have proved:

Any closed surface of genus $g \geq 3$ is a covering surface of a closed surface of genus two.

8-11. *Riemann surfaces made out of equilateral triangles.* Consider a closed topological surface embedded in $\mathbb{R}^3$, say, of genus exceeding one. Triangulate it in any way. Then based on the combinatorics of your triangulation, build a homeomorphic surface R made out of euclidean equilateral triangles. It need not be embedded in $\mathbb{R}^3$. Still, R can be given a complex structure by flattening the vertex angles to 2π. What is remarkable about such a surface are its properties [Jones and Singerman 1996; Schneps 1994; Stephenson 1999; Bowers and Stephenson 2004; Mulase and Penkava 1998]:

The following properties are equivalent:

(i) R is composed of euclidean equilateral triangles.
(ii) There exists a meromorphic function $f : R \to \mathbb{S}^2$ such that its critical values, that is the image of its critical points C, lie in the set $\{0, 1, \infty\}$. (A critical point is a point ζ about which f is not a local homeomorphism; the corresponding critical value is $f(\zeta)$.)
(iii) $R \setminus C$ is a finite cover of the 3-punctured sphere; alternatively, $R \setminus C = \mathbb{H}^2/\Gamma$ where Γ has finite index in the modular group.
(iv) R can be represented by an algebraic curve whose coefficients lie in a finite extension to the field $\mathbb{Q}$ of rational numbers.

The first listed property can be characterized in terms of the existence of a special meromorphic quadratic differential $q\,dz^2$ on R whose critical trajectories divide R into equilateral triangles in the singular euclidean metric associated with $q\,dz^2$. In fact the f-preimage of the segments $[\infty, 0]$, $[0, 1]$, $[1, \infty]$ form a triangular graph on R. Such graphs were called by Grothendieck *dessins d'enfants*, although is seems unlikely a child would come up with one. It has the property that the vertices, the elements of C, can be labeled $+$ or $-$ so that the two endpoints of each edge have the opposite sign. On $\mathbb{S}^2$, the graph complement can be connected and simply connected, so it can just be a finite tree. For a wealth of information about dessins see [Schneps 1994].

These Riemann surfaces, called *Belyĭ surfaces*, can be nicely uniformized by the circle packing technique.

Since there are infinitely many triangulations possible, it is natural to ask whether the Belyĭ surfaces are dense in the Teichmüller space. Compare with Exercise 4-17.

After all, the coefficients of an algebraic curve can be approximated by algebraic numbers.

In some of the following exercises the term "generator pair (A, B)" of a once punctured torus group is used. This means that A, B are loxodromic with parabolic commutator $K = [A, B] = ABA^{-1}B^{-1}$ and generate a discrete group acting on $\mathbb{H}^3$. Thus if A, B are represented by matrices, $\operatorname{tr}(K) = -2$. *We can normalize so that* $K = \begin{pmatrix} -1 & -2 \\ 0 & -1 \end{pmatrix}$.

8-12. Here is a construction, from [Parker and Series 1995], associated with a special generator pair $\langle A, B\rangle$ that gives rise to a convex planar pentagon with four right angles and an ideal vertex.

Let $\boldsymbol{\alpha}^*$ denote the axis of A oriented toward its attracting fixed point, and $\boldsymbol{\beta}^* = B(\boldsymbol{\alpha}^*)$ the axis of $BA^{-1}B^{-1}$ likewise oriented toward its attracting fixed point. We are going to assume that $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$ lie in the same plane P and that P is preserved by A. Such a situation will arise in Exercise 8-20.

Denote the common perpendicular to $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$ by $\boldsymbol{\gamma}$. Find the line $\boldsymbol{\beta}$, perpendicular to $\boldsymbol{\alpha}^*$ such that $A^{-1} = \boldsymbol{\gamma}\boldsymbol{\beta}$. Find the line $\boldsymbol{\alpha}$ perpendicular to $\boldsymbol{\beta}^*$ so that $BA^{-1}B^{-1} = \boldsymbol{\alpha}\boldsymbol{\gamma}$.

For the sixth line $\boldsymbol{\gamma}^*$ we would like to take the axis of $C = A \cdot BA^{-1}B^{-1} = K$. But this is parabolic. So we have instead an ideal line at the ideal vertex which is the fixed point of K.

Now for the side lengths, $\chi(\boldsymbol{\beta}, \boldsymbol{\gamma}) = a^* + \pi i$ and $\chi(\boldsymbol{\gamma}, \boldsymbol{\alpha}) = b^* + \pi i$, but $a^* = b^*$ (why?). Hence, from (8.12), we have

$$\cosh d = \frac{1 + \cosh^2(L/2)}{\sinh^2(L/2)},$$

where d is the distance between the axes of A and BAB^{-1} and L is the translation length of A: $L = 2\log\lambda$ where λ is the larger eigenvalue of A. The transformation A can be assumed to have real eigenvalues since it preserves the plane P. This equation can be transformed by introducing the half-angle formula to become

$$\cosh\frac{d}{2} = \frac{1}{\tanh(L/2)}, \quad \tanh\frac{d}{2} = \frac{1}{\cosh(L/2)}. \tag{8.35}$$

8-13. [Minsky 1999] Suppose (A, B) is a generator pair of a once punctured torus group. Let $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ denote the axes of A, B, $C = AB$ respectively. Prove for the right hexagon $\boldsymbol{\alpha}, \boldsymbol{\gamma}^*, \boldsymbol{\beta}, \boldsymbol{\alpha}^*, \boldsymbol{\gamma}, \boldsymbol{\beta}^*$ that

$$\sinh^2\chi(\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*)\sinh^2\chi(\boldsymbol{\gamma}^*, \boldsymbol{\alpha}^*)\sinh^2\chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) = -1. \tag{8.36}$$

This equation does not depend on how the sides are oriented. *Hint:* To the law of cosines, introduce Lemma 1.5.6(iv) upon recalling Equation (7.12).

8-14. [Minsky 1999] Continuing with the situation of Exercise 8-12, let c^* denote the length of the side on $\boldsymbol{\gamma}^*$.

Consider the ray $\boldsymbol{\beta}^*_o$ of $\boldsymbol{\beta}^*$ that extends from the common endpoint of the sides on $\boldsymbol{\alpha}$, $\boldsymbol{\beta}^*$, runs along the side on $\boldsymbol{\beta}^*$, and ends at $\zeta_1 \in \partial\mathbb{H}^2$. Let ℓ_1 be the ray from ζ_1 that is orthogonal to $\boldsymbol{\gamma}^*$. Let d_1 be the length of segment of the side on $\boldsymbol{\gamma}^*$ cut off by ℓ_1, namely the length of the orthogonal projection of $\boldsymbol{\beta}^*_o$ to $\boldsymbol{\gamma}^*$. Correspondingly take the ray $\boldsymbol{\alpha}^*_o$ of $\boldsymbol{\alpha}^*$ and let d_2 denote the projection of $\boldsymbol{\alpha}^*_o$ onto $\boldsymbol{\gamma}^*$. Prove that

$$c^* \;\le\; d_1 + d_2 + \log 3. \tag{8.37}$$

Hint: We have two quadrilaterals each with three right angles and one zero angle. From Equation (8.14)

$$\sinh^2 \chi(\boldsymbol{\alpha}, \ell_1) \sinh^2 \chi(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) = 1,$$
$$\sinh^2 \chi(\boldsymbol{\beta}, \ell_2) \sinh^2 \chi(\boldsymbol{\gamma}^*, \boldsymbol{\alpha}^*) = 1,$$

and with the help of (8.36) confirm that

$$\sinh^2 \chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) \;=\; -\sinh^2 \chi(\boldsymbol{\alpha}, \ell_1) \sinh^2 \chi(\boldsymbol{\beta}, \ell_2).$$

Now if $d = |\mathrm{Re}\, z|$ then

$$\frac{e^d - 1}{2} \le \frac{e^d - e^{-d}}{2} < |\sinh z| < \frac{e^d + e^{-d}}{2} < e^d.$$

Applying this to the previous equation gives the desired result.

8-15. [Minsky 1999] Given a loxodromic element X define

$$\mathrm{core}_D(X) \;=\; \{\vec{x} \in \mathbb{H}^3 : d(\vec{x}, X\vec{x}) \le L(X) + D\}.$$

Here $L(X)$ is the translation length $2\log|\lambda|$ of X, where λ denotes the larger eigenvalue. Thus $\mathrm{Core}_D(X)$ contains the tube of radius D about the axis of X, but it is larger, especially for small $L(X)$, as X becomes closer to parabolic. In the limit, the core becomes a horoball at the parabolic fixed point.

Prove that for

$$D = 4 \operatorname{arcsinh} 1 + \log 3 = \log 3(1 + \sqrt{2})^4,$$

any generator pair (A, B) of any once-punctured torus group satisfies

$$\mathrm{core}_D(A) \cap \mathrm{core}_D(B) \;\ne\; \varnothing.$$

Hint: Apply Exercise 8-14 to the two quadrilaterals each with three right angles formed by the lines from endpoints of $\boldsymbol{\beta}^*$ and $\boldsymbol{\alpha}^*$ orthogonal to $\boldsymbol{\gamma}^*$. For $p \in c^*$ on the projection of $\boldsymbol{\beta}^*$, the distance of p in particular from the side on $\boldsymbol{\alpha}$ does not exceed $2 \operatorname{arcsinh} 1$ and therefore by triangle inequality $d(p, Ap) \le 4 \operatorname{arcsinh} 1 + L(A)$ (note $A^{-1} = \boldsymbol{\beta}^*\boldsymbol{\gamma}^*$). Likewise $d(p, Bp) \le 4 \operatorname{arcsinh} 1 + L(B)$. By Equation (8.37) the D-cores intersect on the side on $\boldsymbol{\gamma}^*$.

Corollary 8.5.4. *Given a generator pair (A, B) there exists a point $p \in \mathbb{H}^3$ and $\rho < \infty$ such that both $d(p, Ap) < \rho$ and $d(p, Bp) < \rho$.*

8-16. *Hyperbolic Heron's formula.* The following formula for the area $|A|$ of a euclidean triangle is attributed to Heron of Alexandria, who lived almost 2000 years ago. Letting a, b, c denote the side lengths of a euclidean triangle and $s=(a+b+c)/2$ the half-perimeter,

$$|A|^2 = s(s-a)(s-b)(s-c).$$

An analogue of sorts for hyperbolic geometry is as follows [Fenchel 1989]. Let $A, B, C = -(BA)^{-1}$ be normalized matrices. Fix one of the eigenvalues of each of the matrices and denote the choices by $\lambda_A, \lambda_B, \lambda_C$. Set $\chi_A = \log\lambda_A$, $\chi_B = \log\lambda_B$, and $C = \log\lambda_C$. Finally set

$$s = \frac{\chi_A + \chi_B + \chi_C}{2}.$$

The formula states that

$$\operatorname{tr}(ABA^{-1}B^{-1}) - 2 \;=\; 16\cosh s\cosh(s-\chi_A)\cosh(s-\chi_B)\cosh(s-\chi_C). \quad (8.38)$$

Now prove this! (*Hint:* Start by expressing the left hand side first in terms of the traces and then in terms of the eigenvalues.)

Confirm that the numerical value of the right side of (8.38) is invariant under action of the group G of Möbius transformations generated by $\langle A, B\rangle$. That is the generators A, B used in the formula can be replaced by generators $\varphi(A)\,\varphi(B)$ where φ is any automorphism of G. (*Hint:* The commutator $[A, B]$ is independent of φ)

Assume that none of the eigenvalues is ±1 (the traces are not ±2). Show that there exist half-rotation matrices A^*, B^*, C^* with $A = C^*B^*$, $B = A^*C^*$, $C = B^*A^*$ such that the associated axes form a generalized right hexagon, after assigning orientations. Confirm that

$$-\frac{\operatorname{tr}^2(C^*B^*A^*)}{16} \;=\; \cosh s\cosh(s-\chi_A)\cosh(s-\chi_B)\cosh(s-\chi_C).$$

The next four problems outline the development by John Parker and Caroline Series of explicit bending formulas for simple bending in once-punctured torus deformation spaces.

8-17. *The Parker–Series bending formula* [Parker and Series 1995]. Let's start with the following observation. Suppose P is a plane in $\mathbb{H}^3$ and ℓ is a given line in P. Suppose we are also given a Möbius transformation V such that the line $V^{-1}(\ell)$ also lies in P and does not have a common endpoint with ℓ. let $\ell^\perp$ be the line perpendicular to both ℓ and $V^{-1}(\ell)$. Set $\zeta = \ell^\perp \cap V^{-1}(\ell)$. The map V sends P onto a plane $V(P)$ which intersects P along ℓ. It sends ℓ to $V(\ell)$. The line $V(\ell^\perp)$ through $V(\zeta)\in\ell$ is the line perpendicular to ℓ and $V(\ell)$.

Find the midpoint $\zeta_m \in \ell$ between $\ell^\perp\cap\ell$ and $V(\zeta)$. Let ℓ_m be the line through ζ_m, perpendicular to ℓ, such that half-rotation ι_m about ℓ_m sends the line $\ell^\perp$ onto the line $V(\ell^\perp)\subset V(P)$. Necessarily ι_m exchanges the planes P and $V(P)$. Conclude that ι_m also sends $V^{-1}(\ell)$ onto $V(\ell)$.

Now construct the line ℓ_0 orthogonal to P and passing through the midpoint of the segment $[\ell^\perp \cap \ell, \ell^\perp \cap V^{-1}(\ell)]$ of $\ell^\perp$. Let ι_0 denote the half-rotation about ℓ_0. Then

$$\begin{aligned} \iota_m \iota_0(\ell) &= \iota_m(V^{-1}(\ell)) = V(\ell), \\ \iota_m \iota_0(V^{-1}(\ell)) &= \iota_m(\ell) = \ell. \end{aligned}$$

That is, the transformation $X = \iota_m \iota_0$ sends the lines ℓ to $V(\ell)$ and $V^{-1}(\ell)$ to ℓ. Show that this implies that $X = V$. The axis of V must then be the common perpendicular to ℓ_m and ℓ_0.

Next consider $\ell_1 = V(\ell_0) = \iota_m(\ell_0)$ which is orthogonal to $V(\ell^\perp)$ midway between $V(\zeta)$ and $V(\ell)$. Then $\iota_1 = \iota_m \iota_0 \iota_m$ is the corresponding half-rotation and we also have $\iota_1 \iota_m = \iota_m \iota_0 = V$. Thus ℓ_1 is also orthogonal to the axis of V.

With this construction under our belt we can set up the following interesting right hexagon. Suppose U and V are loxodromic transformations whose axes have no common endpoint. Assume the axis $\boldsymbol{\alpha}^*$ of U and $V^{-1}(\boldsymbol{\alpha}^*)$ lie in a plane P. Show that the following six lines determine a right-angled hexagon:

$$\begin{aligned} \boldsymbol{\alpha}^* &= \text{Axis}(U), \\ \boldsymbol{\gamma} &= \text{common perpendicular to } \boldsymbol{\alpha}^* \text{ and } V^{-1}(\boldsymbol{\alpha}^*) = \text{Axis}(V^{-1}UV), \\ \boldsymbol{\beta}^* &= \text{line orthogonal to } P \text{ at midpoint of segment } [\boldsymbol{\gamma} \cap \boldsymbol{\alpha}^*, \boldsymbol{\gamma} \cap V^{-1}(\boldsymbol{\alpha}^*)], \\ \boldsymbol{\alpha} &= \text{Axis}(V), \\ \boldsymbol{\gamma}^* &= V(\boldsymbol{\beta}^*), \\ \boldsymbol{\beta} &= V(\boldsymbol{\gamma}). \end{aligned}$$

To prepare for the law of cosines orient the sides of the hexagon in the usual way. Let $d > 0$ denote the hyperbolic distance between a^* and $V^{-1}(\boldsymbol{\alpha}^*)$. Then

$$\chi(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) = \chi(\boldsymbol{\beta}^* \boldsymbol{\gamma}^*) = \frac{d - i\pi}{2}.$$

Also $\chi(\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*) = \mathcal{L} + \pi i$ where $\mathcal{L} = 2 \log \lambda > 0$ and λ is the larger eigenvalue of V. Therefore

$$\cosh \mathcal{L} = \sinh^2 \frac{d}{2} + \cosh^2 \frac{d}{2} \cosh \chi(\boldsymbol{\beta}, \boldsymbol{\gamma}).$$

Specialize to the case that (A, B) is a generator pair for a once-punctured torus group; this means that A, B are loxodromic generators with parabolic commutator $ABA^{-1}B^{-1}$. Set $U = A$ and $V = B$ so that in the formula above d is the distance between the axes of A and $B^{-1}AB$. Assume that A preserves the plane P so the eigenvalues of A are real; L_A will denote the larger eigenvalue. Let $\mathcal{L}_B$ denote the larger eigenvalue of B; $\mathcal{L}_B$ will not in general be real so the translation length of B is $|\mathcal{L}_B|$. However the real part $\mathcal{L}_B$ is positive. Now we can incorporate (8.35): apply the double angle formula to $\cosh \mathcal{L}_B$ and substitute for $\cosh \frac{d}{2}$ and $\sinh \frac{d}{2}$ to end up with the beautiful formula:

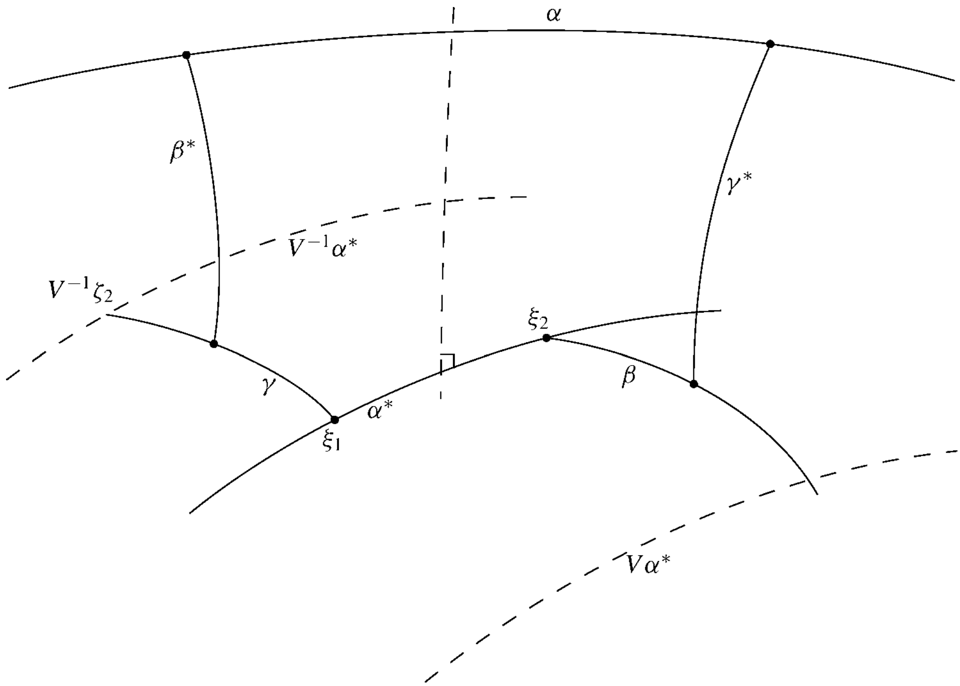


Fig. 8.10. A Parker–Series right hexagon expressing bending.

The Parker–Series Bending Formula 8.5.5.

$$\cosh^2 \frac{\chi(\boldsymbol{\beta}, \boldsymbol{\gamma})}{2} = \cosh^2 \frac{L_A}{2} \tanh^2 \frac{\mathcal{L}_B}{2}. \tag{8.39}$$

We will digress in order to interpret $\chi(\boldsymbol{\beta}, \boldsymbol{\gamma})$. Consider the vertical half plane Q based on $\mathbb{R}$ in the upper half space model of $\mathbb{H}^3$. Suppose a fuchsian once-punctured torus group is acting in Q. Let $\boldsymbol{\alpha}^*$ denote the positive vertical axis oriented toward ∞ and assume it is the axis a hyperbolic generator X. Now X corresponds to a simple loop α, not retractable to the puncture, on the quotient punctured torus and $\boldsymbol{\alpha}^*$ is a lift of α. Choose a simple loop σ which crosses α exactly once, and from left to right. Fix a lift σ^* which crosses the line $\boldsymbol{\alpha}^*$ from its left side to its right at a point ζ and is the axis of a Möbius transformation Y. We may arrange things so that (X, Y) is a generator pair.

We will deform the fuchsian group $\langle X, Y\rangle$ by a *quakebend*, also known as a *complex earthquake*, or *shearing and bending*. We start by giving a complex number of the form $\kappa = \log\rho + \phi i$, with $\rho > 0$ and $-\pi < \phi < \pi$.

Apply the transformation $S_\rho : z = x + iy \mapsto \rho z$ to the *right* half of P. In particular the point $\zeta \in \boldsymbol{\alpha}^*$ is moved up or down along $\boldsymbol{\alpha}^*$, depending on whether ρ is > 1 or < 1, signed hyperbolic distance $\log\rho$. Correspondingly the right half of the line σ^* is moved up or down and becomes a half line from $\rho\zeta$. The point $B(\zeta)$ on the right half of σ moves to $\rho Y(\zeta) \in P$. Actually S is the restriction of Möbius transformation of $\mathbb{H}^3$.

Next rotate the right half of Q about the vertical $\boldsymbol{\alpha}^*$ by angle ϕ, measured so that $\phi = 0$ corresponds to no bending. Denote this elliptic transformation by E_ϕ. The point $E_\phi Y(\zeta)$ rotates off P to a point $\hat{\zeta} \in \mathbb{H}^3$.

This process results in the *deformation* $Y \mapsto Y_\phi = E_\phi S_\rho Y$ and

$$G = \langle X, Y \rangle \mapsto G_\phi = \langle X, Y_\phi \rangle.$$

This will be a new punctured torus group provided κ is sufficiently small and that the deformed generators still satisfy the trace relation. Note that the quakebend depends on two real parameters; preservation of the commutator trace -2 gives rise to two real equations.

For the group G_ϕ, the lines $\boldsymbol{\alpha}^*$ and $Y_\phi(\boldsymbol{\alpha}^*)$ lie in the same plane Q_ϕ — the rotated right half of Q. The angle ϕ is the dihedral angle between ${Y_\phi}^{-1}(Q_\phi)$ and Q_ϕ.

Return to the situation preceding (8.39). Let X correspond to A and Y to B. In the Parker–Series equation (8.39),

$$\chi(\boldsymbol{\beta}, \boldsymbol{\gamma}) \;=\; \kappa \;\log \rho + \phi i,$$

where $\log \rho > 0$ is the distance along $\boldsymbol{\alpha}^*$ from $\boldsymbol{\alpha}^* \cap \boldsymbol{\beta}$ to $\boldsymbol{\alpha}^* \cap \boldsymbol{\gamma}$, and ϕ is the dihedral angle from P to $B(P)$.

Bend the right half-plane abutting $\boldsymbol{\alpha}^*$ with respect to the left half-plane so that the two make the dihedral angle ϕ where $\phi = 0$ corresponds to no bending at all, $-\pi < \phi < \pi$. Let $x \in \mathbb{R}$ be a given number. The line σ is broken in two parts. The left part ends at $\zeta \in \boldsymbol{\alpha}^*$. From ζ continue for x units along $\boldsymbol{\alpha}^*$ (in the positive or negative direction depending on the sign of x) reaching a point $\zeta_1 \in \boldsymbol{\alpha}^*$. Now continue in the right half plane from ζ_1, making the same angle with $\boldsymbol{\alpha}^*$ as the original σ.

Corollary 8.5.6. *Suppose the sequence of generator pairs* $\{(A_n, B_n)\}$ *are the result of bending along the axis of* A_n. *Assume that* $\lim A_n$ *is a parabolic transformation. Then on a subsequence either* $\lim |\mathcal{L}_{B_n}| = \infty$, *or* $\lim \rho_n = 1$ *and* $\lim \phi_n = \pi$.

Can the second possibility occur?

8-18. *Real traces.*Suppose (A, B) is the generator pair of a fuchsian once punctured torus group. Assume that the axis of B is orthogonal to the axis of A. Bend B along the axis of A getting a new element B_ϕ. Show that the trace of B_ϕ is real.

Hint: Assume A, B act in the vertical half plane along $\mathbb{R}$ in the upper half space model and the axis of A is the vertical axis from the origin. Show that $B_\phi = E_\phi B$, where

$$E_\phi = \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix}.$$

8-19. [Parker and Series 1995]Choose loxodromic Möbius transformations A and B so that the common perpendicular to their axes is the vertical half-line from $0 \in \mathbb{C}$ in the upper half space model of $\mathbb{H}^3$. Show that their fixed points are necessarily symmetric with respect to 0. Normalize A to have fixed points ± 1. Write the fixed points of B as $\pm re^{i\theta}$. Show that if both A and B have real traces, which we may take to be ≥ 0, then

$$A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}, \quad B = \begin{pmatrix} u & vre^{i\theta} \\ (v/r)\,e^{-i\theta} & u \end{pmatrix},$$
$$a^2 - b^2 = u^2 - v^2 = 1, \quad a, b, u, v > 0.$$

We will require as well that $\operatorname{tr}[A, B] = -2$. Using Proposition 7.2.1, show that necessarily $\theta \neq 0$ and the trace condition is satisfied if and only if either $r = 1$ or $\theta = \pi/2$. We will work only with the case that $\theta = \pi/2$ which holds if and only if

$$bv\left(r + \frac{1}{r}\right) = 2, \quad b, v > 0, \ r \geq 1. \tag{8.40}$$

Then the group $G = \langle A, B \rangle$ depends only on the two real numbers $b, v > 0$. When $r = 1$, that is when $bv \sin\theta = 1$, show that G is fuchsian, preserving the unit disk.

Show that each of $\mathfrak{I}(A)$ and $\mathfrak{I}(A^{-1})$ is tangent to $\mathfrak{I}(B)$ and to $\mathfrak{I}(B^{-1})$ at a point on the line segment between the centers. Therefore the common exterior of the four circles has two components, an inner one containing 0 and an outer one containing ∞. Each circular polygon is bounded by four sides. Show that A and B pair the opposite sides of both components (see §1.6). Conclude that the points of tangency are parabolic fixed points of $[A, B]$ and its conjugates. Therefore the quotient consists of two once punctured tori and (A, B) is a generator pair of a group in our space. In terms of complex probabilities introduced in Exercise 1-37, what are their coordinates? Then draw the fundamental region for some parameter values using Wada's computer program.

The common exterior of the planes determined by the four circles forms the Ford fundamental polyhedron, which therefore has only four faces.

8-20. *The single bend formula* [Parker and Series 1995]. Carry on with the two real parameter class of once punctured torus groups $G = \langle A, B \rangle$ introduced in Exercise 8-19 above. These represent a two dimensional slice through the four real dimensional deformation space $\mathfrak{T}$. Assume that $r > 1$ so that G is not fuchsian. We are going to explicitly describe how each of the two boundary components of the convex hull are bent. It will turn out that each component arises by bending along a single line and then its conjugates: the groups are obtained by pure bending along the axis of A, and automatically, along the axis of B. The two bending angles serve also as parameters for the class of groups G.

Start by establishing the symmetry with respect to the reflections J in the real axis and $J_\perp$ in the imaginary axis: $JAJ = A$, $J_\perp A J_\perp = A^{-1}$, and $JBJ = B^{-1}$, $J_\perp B J_\perp = B$. The symmetries preserve $\langle A, B \rangle$ as well. Therefore the limit set and the convex hull are also invariant under these symmetries.

Using symmetry show that $\Lambda(G)$ intersects the real and imaginary axes only in the four fixed points $\{\pm 1, \pm ri\}$. This sets the stage for showing the axes of A and B lie in the opposite boundary components of the convex hull.

Let $\mathcal{C}$ denote the convex hull of the limit set as in §3.10. First examine the vertical half plane P resting on $\mathbb{R}$. This plane contains the axis of A, and the axis $\boldsymbol{\alpha}$ of A is automatically contained in $\mathcal{C}$. If $\boldsymbol{\alpha}$ is not on the boundary $\partial\mathcal{C}$ then $P \cap \partial\mathcal{C}$ consists of two convex arcs σ_1, σ_2 which are separated in P by $\boldsymbol{\alpha}$, one on each boundary component Σ_1, Σ_2 of $\mathcal{C}$. By the paragraph above, the end points of these arcs must be the fixed points of A. If either is a geodesic it must agree with $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha} \subset \partial\mathcal{C}$. At most one can be $\boldsymbol{\alpha}$. Each of σ_1, σ_2 which is not a geodesic has bends in it. For such to occur, there must be a (geodesic) line or lines $\boldsymbol{l}_1, \boldsymbol{l}_2$ in the same boundary component or components of $\mathcal{C}$ and which intersect P.

Now $J(\boldsymbol{l}_1)$, say, must also be a line in Σ_1 yet it would cross $\boldsymbol{l}_1$ unless $J(\boldsymbol{l}_1) = \boldsymbol{l}_1$. Hence $\boldsymbol{l}_1$ is orthogonal to the plane P. Show that $\boldsymbol{l}_1$ must in fact lie in the vertical plane $P_\perp$ resting on the imaginary axis. Once we know this, we can conclude that $\boldsymbol{l}_1 = \boldsymbol{\beta}$. Since this cannot also hold for $\boldsymbol{l}_2$ it follows that $\boldsymbol{l}_1 = \boldsymbol{\alpha} \subset \Sigma_1$. Since the argument applies to *all* bending lines $\boldsymbol{l}_1 \subset \Sigma_1$ that are transverse to P it shows that $\boldsymbol{\beta}$ is the only such line.

Repeating the argument deduce that $\boldsymbol{\beta} \subset \Sigma_2$ and $\boldsymbol{\alpha}$ is the only bending line that is transverse to $P_\perp$.

Summing up, the axis of A lies in Σ_1. All its conjugates in $G = \langle A, B\rangle$ do as well since Σ_1 is invariant under G. These are the totality of the bending lines on Σ_1, and all the bending angles are the same. Likewise the axis of B and its conjugates comprise the bending lines of Σ_2, all with the same bending angles. The conditions Equation (8.40) on the 2-real parameter groups $\langle A, B\rangle$ are *necessary and sufficient that the corresponding convex hull boundary components are bent along the axes of A and B.*

The endpoints of the bending lines on the two components separate each other on $\Lambda(G)$.

Apply to this case the Parker–Series bending formula. We first point out that the assumption made in Exercise 8-12 holds, namely that the axis $\boldsymbol{\alpha}$ of A and the axis $B^{-1}(\boldsymbol{\alpha})$ of $B^{-1}AB$ lie in the same plane: For the bending lines in Σ_1 separate it into infinitely many components. Each component is contained in the hyperbolic plane determined by any two of its infinitely many boundary components.

Return to Equation (8.39). In the present case the translation lengths $L_A, L_B = \mathcal{L}_B$ have been taken to be both real and positive. The left side must be positive as well. Since $\chi(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \log\rho + i\phi$, this is possible only if $\rho = 1$ so that the group G is obtained by pure bending. Therefore for Σ_1,

$$\cos\frac{\phi_A}{2} = \cosh\frac{L_A}{2}\tanh\frac{L_B}{2}.$$

Interchanging the roles of A and B for Σ_2,

$$\cos\frac{\phi_B}{2} = \cosh\frac{L_B}{2}\tanh\frac{L_A}{2}.$$

Here $\phi_A, \phi_B \in (0, \pi)$ are the bending angles at the axes of A and B. In terms of the matrix A, L_A has the expression

$$L_A = 2\log(a+b), \quad \cosh\frac{L_A}{2} = a, \quad \sinh\frac{L_A}{2} = b,$$

and the analogous equations in terms of u, v hold for L_B. Show that $\phi_A = 0$ if and only if $bv = 1$, that is, $r = 1$ and G is fuchsian. In this case also $\phi_B = 0$.

Solving for L_A, L_B when $r > 1$, we end up with

The Parker–Series Single Bend Formula 8.5.7.

$$b = \sinh\frac{L_A}{2} = \sin\frac{\phi_A}{2}\cot\frac{\phi_B}{2}, \quad v = \sinh\frac{L_B}{2} = \sin\frac{\phi_B}{2}\cot\frac{\phi_A}{2}. \tag{8.41}$$

In particular, there is a homeomorphism between pairs of angles $\phi_A, \phi_B \in (0, \pi)$ and (nonfuchsian) quasifuchsian groups with convex hull boundaries bent along the axes of A, B at angles ϕ_A, ϕ_B respectively.

Bibliography

[Abikoff and Maskit 1977] W. Abikoff and B. Maskit, "Geometric decompositions of Kleinian groups", *Amer. J. Math.* **99**:4 (1977), 687–697.

[Accola 1966] R. D. M. Accola, "Invariant domains for Kleinian groups", *Amer. J. Math.* **88** (1966), 329–336.

[Adams 1987] C. C. Adams, "The noncompact hyperbolic 3-manifold of minimal volume", *Proc. Amer. Math. Soc.* **100**:4 (1987), 601–606.

[Agol 2002] I. Agol, "Volume change under drilling", *Geom. Topol.* **6** (2002), 905–916.

[Agol 2004] I. Agol, "Tameness of hyperbolic 3-manfolds", preprint, 2004. Available at www.arXiv.org/abs/math.GT/0405568.

[Agol et al. 2005] I. Agol, P. Storm, and W. P. Thurston, "Lower bounds on volumes of hyperbolic Haken 3-manifolds", preprint, 2005. Available at arXiv:math.DG/0506338. With an appendix by N. Dunfield.

[Ahlfors 1964] L. V. Ahlfors, "Finitely generated Kleinian groups", *Amer. J. Math.* **86** (1964), 413–429.

[Ahlfors 1966] L. V. Ahlfors, *Lectures on quasiconformal mappings*, Van Nostrand Mathematical Studies **10**, Van Nostrand, Toronto, 1966.

[Ahlfors 1973] L. V. Ahlfors, *Conformal invariants: topics in geometric function theory*, McGraw-Hill, New York, 1973.

[Ahlfors 1978] L. V. Ahlfors, *Complex analysis*, McGraw-Hill, New York, 1978.

[Ahlfors 1981] L. V. Ahlfors, *Möbius transformations in several dimensions*, Ordway Professorship Lectures in Mathematics, Univ. of Minnesota, Minneapolis, MN, 1981.

[Ahlfors and Sario 1960] L. V. Ahlfors and L. Sario, *Riemann surfaces*, Princeton Mathematical Series **26**, Princeton Univ. Press, Princeton, NJ, 1960.

[Akiyoshi et al. 1999] H. Akiyoshi, M. Sakuma, M. Wada, and Y. Yamashita, "Ford domains of punctured torus groups and two-bridge knot groups", in *Knot Theory* (Proceedings of the workshop held in Toronto dedicated to 70th birthday of Prof. K. Murasugi), 1999.

[Akiyoshi et al. 2003] H. Akiyoshi, M. Sakuma, M. Wada, and Y. Yamashita, "Jørgensen's picture of punctured torus groups and its refinement", pp. 247–273 in *Kleinian groups and hyperbolic 3-manifolds* (Warwick, 2001), edited by Y. Komori et al., London Math. Soc. Lecture Note Ser. **299**, Cambridge Univ. Press, Cambridge, 2003.

[Akiyoshi et al. 2005] H. Akiyoshi, M. Sakuma, M. Wada, and Y. Yamashita, "Punctured torus groups and 2-bridge knot groups", preprint, 2005.

[Anderson 1994] J. W. Anderson, "Intersections of analytically and geometrically finite subgroups of Kleinian groups", *Trans. Amer. Math. Soc.* **343**:1 (1994), 87–98.

[Anderson 1996] J. W. Anderson, "The limit set intersection theorem for finitely generated Kleinian groups", *Math. Res. Lett.* **3**:5 (1996), 675–692.

[Anderson and Canary 1996a] J. W. Anderson and R. D. Canary, "Algebraic limits of Kleinian groups which rearrange the pages of a book", *Invent. Math.* **126**:2 (1996), 205–214.

[Anderson and Canary 1996b] J. W. Anderson and R. D. Canary, "Cores of hyperbolic 3-manifolds and limits of Kleinian groups", *Amer. J. Math.* **118**:4 (1996), 745–779.

[Anderson et al. 1996] J. W. Anderson, R. D. Canary, M. Culler, and P. B. Shalen, "Free Kleinian groups and volumes of hyperbolic 3-manifolds", *J. Differential Geom.* **43**:4 (1996), 738–782.

[Anderson et al. 2000] J. W. Anderson, R. D. Canary, and D. McCullough, "The topology of deformation spaces of Kleinian groups", *Ann. of Math.* (2) **152**:3 (2000), 693–741.

[Ballmann et al. 1985] W. Ballmann, M. Gromov, and V. Schroeder, *Manifolds of nonpositive curvature*, Progress in Mathematics **61**, Birkhäuser, Boston, 1985.

[Bass 1980] H. Bass, "Groups of integral representation type", *Pacific J. Math.* **86**:1 (1980), 15–51.

[Beardon 1983] A. F. Beardon, *The geometry of discrete groups*, Graduate Texts in Mathematics **91**, Springer, New York, 1983.

[Beardon and Jørgensen 1975] A. F. Beardon and T. Jørgensen, "Fundamental domains for finitely generated Kleinian groups", *Math. Scand.* **36** (1975), 21–26.

[Beardon and Maskit 1974] A. F. Beardon and B. Maskit, "Limit points of Kleinian groups and finite sided fundamental polyhedra", *Acta Math.* **132** (1974), 1–12.

[Beardon and Pommerenke 1978] A. F. Beardon and C. Pommerenke, "The Poincaré metric of plane domains", *J. London Math. Soc.* (2) **18**:3 (1978), 475–483.

[Beardon and Stephenson 1990] A. F. Beardon and K. Stephenson, "The uniformization theorem for circle packings", *Indiana Univ. Math. J.* **39**:4 (1990), 1383–1425.

[Benedetti and Petronio 1992] R. Benedetti and C. Petronio, *Lectures on hyperbolic geometry*, Universitext, Springer, Berlin, 1992.

[Bers 1970a] L. Bers, "On boundaries of Teichmüller spaces and on Kleinian groups, I", *Ann. of Math.* (2) **91** (1970), 570–600.

[Bers 1970b] L. Bers, "Spaces of Kleinian groups", pp. 9–34 in *Several Complex Variables* (College Park, MD, 1970), vol. I, edited by J. Horváth, Lecture Notes in Math. **155**, Springer, Berlin, 1970.

[Bers 1985] L. Bers, "An inequality for Riemann surfaces", pp. 87–93 in *Differential geometry and complex analysis*, Springer, Berlin, 1985.

[Bers and Greenberg 1971] L. Bers and L. Greenberg, "Isomorphisms between Teichmüller spaces", pp. 53–79 in *Advances in the Theory of Riemann Surfaces* (Stony Brook, NY, 1969), edited by L. V. Ahlfors et al., Ann. of Math. Studies **66**, Princeton Univ. Press, Princeton, N.J., 1971.

[Besson et al. 1999] G. Besson, G. Courtois, and S. Gallot, "Lemme de Schwarz réel et applications géométriques", *Acta Math.* **183**:2 (1999), 145–169.

[Bestvina 1988] M. Bestvina, "Degenerations of the hyperbolic space", *Duke Math. J.* **56**:1 (1988), 143–161.

[Birman 1974] J. S. Birman, "Mapping class groups of surfaces: a survey", pp. 57–71 in *Discontinuous groups and Riemann surfaces* (College Park, MD, 1973), edited by L. Greenberg, Ann. of Math. Studies **79**, Princeton Univ. Press, Princeton, NJ, 1974.

[Birman and Series 1985] J. S. Birman and C. Series, "Geodesics with bounded intersection number on surfaces are sparsely distributed", *Topology* **24**:2 (1985), 217–225.

[Bishop and Jones 1997] C. J. Bishop and P. W. Jones, "Hausdorff dimension and Kleinian groups", *Acta Math.* **179**:1 (1997), 1–39.

[Bleiler and Hodgson 1996] S. A. Bleiler and C. D. Hodgson, "Spherical space forms and Dehn filling", *Topology* **35**:3 (1996), 809–833.

[Bobenko and Springborn 2004] A. I. Bobenko and B. A. Springborn, "Variational principles for circle patterns and Koebe's theorem", *Trans. Amer. Math. Soc.* **356**:2 (2004), 659–689.

[Boileau and Porti 2001] M. Boileau and J. Porti, *Geometrization of 3-orbifolds of cyclic type*, Astérisque **272**, 2001.

[Boileau et al. 2003] M. Boileau, S. Maillot, and J. Porti, *Three-dimensional orbifolds and their geometric structures*, Panoramas et Synthèses **15**, Société Mathématique de France, Paris, 2003.

[Boileau et al. 2005] M. Boileau, B. Leeb, and J. Porti, "Geometrization of 3-dimensional orbifolds", *Ann. of Math.* (2) **162**:1 (2005), 195–290.

[Bonahon 1986] F. Bonahon, "Bouts des variétés hyperboliques de dimension 3", *Ann. of Math.* (2) **124**:1 (1986), 71–158.

[Bonahon 1988] F. Bonahon, "The geometry of Teichmüller space via geodesic currents", *Invent. Math.* **92**:1 (1988), 139–162.

[Bonahon 1996] F. Bonahon, "Shearing hyperbolic surfaces, bending pleated surfaces and Thurston's symplectic form", *Ann. Fac. Sci. Toulouse Math.* (6) **5**:2 (1996), 233–297.

[Bonahon 2001] F. Bonahon, "Geodesic laminations on surfaces", pp. 1–37 in *Laminations and foliations in dynamics, geometry and topology* (Stony Brook, NY, 1998), edited by M. Lyubich et al., Contemp. Math. **269**, Amer. Math. Soc., Providence, RI, 2001.

[Bonahon 2002] F. Bonahon, "Geometric structures on 3-manifolds", pp. 93–164 in *Handbook of geometric topology*, edited by R. J. Daverman and R. B. Sher, North-Holland, Amsterdam, 2002.

[Bonahon and Otal 2004] F. Bonahon and J.-P. Otal, "Laminations mesurées de plissage des variétés hyperboliques de dimension 3", *Ann. Math.* **160** (2004), 1013–1055.

[Bonk 2002] M. Bonk, "Singular surfaces and meromorphic functions", *Notices Amer. Math. Soc.* **49**:6 (2002), 647–657.

[Bowditch 1996] B. H. Bowditch, "A proof of McShane's identity via Markoff triples", *Bull. London Math. Soc.* **28**:1 (1996), 73–78.

[Bowditch 1998] B. H. Bowditch, "Markoff triples and quasi-Fuchsian groups", *Proc. London Math. Soc.* (3) **77**:3 (1998), 697–736.

[Bowers and Stephenson 1992] P. L. Bowers and K. Stephenson, "The set of circle packing points in the Teichmüller space of a surface of finite conformal type is dense", *Math. Proc. Cambridge Philos. Soc.* **111**:3 (1992), 487–513.

[Bowers and Stephenson 2004] P. L. Bowers and K. Stephenson, *Uniformizing dessins and Belyĭ maps via circle packing*, Mem. Amer. Math. Soc. **805**, Amer. Math. Soc., Providence, 2004.

[Brendle and Farb 2004] T. E. Brendle and B. Farb, "Every mapping class group is generated by 6 involutions", *J. Algebra* **278**:1 (2004), 187–198.

[Bridgeman 1998] M. Bridgeman, "Average bending of convex pleated planes in hyperbolic three-space", *Invent. Math.* **132**:2 (1998), 381–391.

[Bridgeman 2003] M. Bridgeman, "Bounds on the average bending of the convex hull boundary of a Kleinian group", *Michigan Math. J.* **51**:2 (2003), 363–378.

[Brock 2000] J. F. Brock, "Continuity of Thurston's length function", *Geom. Funct. Anal.* **10**:4 (2000), 741–797.

[Brock 2001a] J. F. Brock, "Boundaries of Teichmüller spaces and end-invariants for hyperbolic 3-manifolds", *Duke Math. J.* **106**:3 (2001), 527–552.

[Brock 2001b] J. F. Brock, "Iteration of mapping classes and limits of hyperbolic 3-manifolds", *Invent. Math.* **143**:3 (2001), 523–570.

[Brock and Bromberg 2003] J. F. Brock and K. W. Bromberg, "Cone-manifolds and the density conjecture", pp. 75–93 in *Kleinian groups and hyperbolic 3-manifolds (Warwick, 2001)*, London Math. Soc. Lecture Note Ser. **299**, Cambridge Univ. Press, Cambridge, 2003.

[Brock and Bromberg 2004] J. F. Brock and K. W. Bromberg, "On the density of geometrically finite Kleinian groups", *Acta Math.* **192**:1 (2004), 33–93.

[Brock and Souto 2006] J. Brock and J. Souto, "Algebraic limits of geometrically finite manifolds are tame", *Geom. Funct. Anal.* (2006). to appear.

[Brock et al. 2003] J. Brock, K. Bromberg, R. Evans, and J. Souto, "Tameness on the boundary and Ahlfors' measure conjecture", *Publ. Math. Inst. Hautes Études Sci.* no. 98 (2003), 145–166.

[Brock et al. 2004] J. Brock, R. D. Canary, and Y. Minsky, "Classification of Kleinian surface groups, II: The ending lamination conjecture", preprint, 2004. Available at www.arXiv.org/abs/math.GT/0412006.

[Brock et al. $\geq$ 2007] J. Brock, R. D. Canary, and Y. Minsky, "The classification of finitely-generated Kleinian groups", Technical report. to appear.

[Bromberg 2000] K. Bromberg, "Hyperbolic Dehn surgery on geometrically infinite 3-manifolds", preprint, 2000. Available at www.arXiv.org/math.GT/0009150.

[Bromberg 2004] K. Bromberg, "Hyperbolic cone-manifolds, short geodesics, and Schwarzian derivatives", *J. Amer. Math. Soc.* **17**:4 (2004), 783–826.

[Bromberg 2006] K. Bromberg, "The space of Kleinian punctured torus groups is not locally connected", preprint, 2006. Available at http://www.math.utah.edu/~bromberg/.

[Bromberg and Holt 2001] K. Bromberg and J. Holt, "Self-bumping of deformation spaces of hyperbolic 3-manifolds", *J. Differential Geom.* **57**:1 (2001), 47–65.

[Brooks 1985] R. Brooks, "On the deformation theory of classical Schottky groups", *Duke Math. J.* **52**:4 (1985), 1009–1024.

[Brooks 1986] R. Brooks, "Circle packings and co-compact extensions of Kleinian groups", *Invent. Math.* **86**:3 (1986), 461–469.

[Brooks 1999] R. Brooks, "Twist surfaces", pp. 85–103 in *Random walks and discrete potential theory* (Cortona, 1997), edited by M. Picardello and W. Woess, Symposia Mathematica **39**, Cambridge Univ. Press, Cambridge, 1999.

[Brooks and Matelski 1981] R. Brooks and J. P. Matelski, "The dynamics of 2-generator subgroups of PSL(2, $\mathbb{C}$)", pp. 65–71 in *Riemann surfaces and related topics* (Stony Brook, NY, 1978), edited by I. Kra and B. Maskit, Ann. of Math. Stud. **97**, Princeton Univ. Press, Princeton, NJ, 1981.

[Brunner et al. 1984] A. M. Brunner, M. L. Frame, Y. W. Lee, and N. J. Wielenberg, "Classifying torsion-free subgroups of the Picard group", *Trans. Amer. Math. Soc.* **282**:1 (1984), 205–235.

[Burger and Canary 1994] M. Burger and R. D. Canary, "A lower bound on λ_0 for geometrically finite hyperbolic n-manifolds", *J. Reine Angew. Math.* **454** (1994), 37–57.

[Burger et al. 2002] M. Burger, T. Gelander, A. Lubotzky, and S. Mozes, "Counting hyperbolic manifolds", *Geom. Funct. Anal.* **12**:6 (2002), 1161–1173.

[Buser 1992] P. Buser, *Geometry and spectra of compact Riemann surfaces*, Progress in Mathematics **106**, Birkhäuser, Boston, 1992.

[Button 2005] J. O. Button, "Fibred and virtually fibred hyperbolic 3-manifolds in the censuses", *Experiment. Math.* **14**:2 (2005), 231–255.

[Calegari and Gabai 2004] D. Calegari and D. Gabai, "Shrinkwrapping and the taming of hyperbolic 3-manifolds", preprint, 2004. Available at www.arXiv.org/abs/math.GT/0407161.

[Callahan et al. 1999] P. J. Callahan, J. C. Dean, and J. R. Weeks, "The simplest hyperbolic knots", *J. Knot Theory Ramifications* **8**:3 (1999), 279–297.

[Canary 1991] R. D. Canary, "The Poincaré metric and a conformal version of a theorem of Thurston", *Duke Math. J.* **64**:2 (1991), 349–359.

[Canary 1992] R. D. Canary, "On the Laplacian and the geometry of hyperbolic 3-manifolds", *J. Differential Geom.* **36**:2 (1992), 349–367.

[Canary 1993] R. D. Canary, "Ends of hyperbolic 3-manifolds", *J. Amer. Math. Soc.* **6**:1 (1993), 1–35.

[Canary 1996] R. D. Canary, "A covering theorem for hyperbolic 3-manifolds and its applications", *Topology* **35**:3 (1996), 751–778.

[Canary and Hersonsky 2004] R. D. Canary and S. Hersonsky, "Ubiquity of geometric finiteness in boundaries of deformation spaces of hyperbolic 3-manifolds", *Amer. J. Math.* **126**:6 (2004), 1193–1220.

[Canary and McCullough 2004] R. D. Canary and D. McCullough, *Homotopy equivalences of 3-manifolds and deformation theory of Kleinian groups*, Mem. Amer. Math. Soc. **812**, Amer. Math. Soc., Providence, 2004.

[Canary and Minsky 1996] R. D. Canary and Y. N. Minsky, "On limits of tame hyperbolic 3-manifolds", *J. Differential Geom.* **43**:1 (1996), 1–41.

[Canary and Taylor 1994] R. D. Canary and E. Taylor, "Kleinian groups with small limit sets", *Duke Math. J.* **73**:2 (1994), 371–381.

[Canary et al. 1987] R. D. Canary, D. B. A. Epstein, and P. Green, "Notes on notes of Thurston", pp. 3–92 in *Analytical and geometric aspects of hyperbolic space* (Warwick and Durham, 1984), edited by D. B. A. Epstein, London Math. Soc. Lecture Note Ser. **111**, Cambridge Univ. Press, Cambridge, 1987. Reprinted in [Canary et al. 2006].

[Canary et al. 1999] R. D. Canary, Y. N. Minsky, and E. C. Taylor, "Spectral theory, Hausdorff dimension and the topology of hyperbolic 3-manifolds", *J. Geom. Anal.* **9**:1 (1999), 17–40.

[Canary et al. 2003] R. D. Canary, M. Culler, S. Hersonsky, and P. B. Shalen, "Approximation by maximal cusps in boundaries of deformation spaces of Kleinian groups", *J. Differential Geom.* **64**:1 (2003), 57–109.

[Canary et al. 2006] R. D. Canary, A. Marden, and D. B. A. Epstein (editors), *Fundamentals of hyperbolic geometry: selected expositions*, edited by R. D. Canary et al., London Math. Soc. Lecture Note Ser. **328**, Cambridge Univ. Press, Cambridge, 2006.

[Cannon 1991] J. W. Cannon, "The theory of negatively curved spaces and groups", pp. 315–369 in *Ergodic theory, symbolic dynamics, and hyperbolic spaces (Trieste, 1989)*, Oxford Sci. Publ., Oxford Univ. Press, New York, 1991.

[Cannon 2002] J. W. Cannon, "Geometric group theory", pp. 261–305 in *Handbook of geometric topology*, North-Holland, Amsterdam, 2002.

[Cannon and Cooper 1992] J. W. Cannon and D. Cooper, "A characterization of cocompact hyperbolic and finite-volume hyperbolic groups in dimension three", *Trans. Amer. Math. Soc.* **330**:1 (1992), 419–431.

[Cannon and Thurston 1989] J. W. Cannon and W. Thurston, "Group invariant Peano curves", preprint, 1989.

[Cannon et al. 1997] J. W. Cannon, W. J. Floyd, R. Kenyon, and W. R. Parry, "Hyperbolic geometry", pp. 59–115 in *Flavors of geometry*, Math. Sci. Res. Inst. Publ. **31**, Cambridge Univ. Press, Cambridge, 1997.

[Cao 1994] J. G. Cao, "The Bers-Nielsen kernels and souls of open surfaces with negative curvature", *Michigan Math. J.* **41**:1 (1994), 13–30.

[Cao and Meyerhoff 2001] C. Cao and G. R. Meyerhoff, "The orientable cusped hyperbolic 3-manifolds of minimum volume", *Invent. Math.* **146**:3 (2001), 451–478.

[Cao and Zhu 2006] H.-D. Cao and 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** (2006), 165–492.

[Casson and Bleiler 1988] A. J. Casson and S. A. Bleiler, *Automorphisms of surfaces after Nielsen and Thurston*, London Mathematical Society Student Texts **9**, Cambridge Univ. Press, Cambridge, 1988.

[Casson and Jungreis 1994] A. Casson and D. Jungreis, "Convergence groups and Seifert fibered 3-manifolds", *Invent. Math.* **118**:3 (1994), 441–456.

[Chavel 1993] I. Chavel, *Riemannian geometry—a modern introduction*, Cambridge Tracts in Mathematics **108**, Cambridge Univ. Press, Cambridge, 1993.

[Choi and Series 2006] Y.-E. Choi and C. Series, "Lengths are coordinates for convex structures", *J. Differential Geom.* **73**:1 (2006), 75–117.

[Chow and Knopf 2004] B. Chow and D. Knopf, *The Ricci flow: an introduction*, Mathematical Surveys and Monographs **110**, Amer. Math. Soc., Providence, RI, 2004.

[Conder et al. 2005] M. Conder, G. Martin, and A. Torstensson, "Maximal symmetry groups of hyperbolic 3-manifolds", preprint, 2005.

[Cooper 1999] D. Cooper, "The volume of a closed hyperbolic 3-manifold is bounded by π times the length of any presentation of its fundamental group", *Proc. Amer. Math. Soc.* **127**:3 (1999), 941–942.

[Cooper and Lackenby 1998] D. Cooper and M. Lackenby, "Dehn surgery and negatively curved 3-manifolds", *J. Differential Geom.* **50**:3 (1998), 591–624.

[Cooper et al. 2000] D. Cooper, C. D. Hodgson, and S. P. Kerckhoff, *Three-dimensional* *orbifolds and cone-manifolds*, MSJ Memoirs **5**, Mathematical Society of Japan, Tokyo, 2000.

[Coulson et al. 2000] D. Coulson, O. A. Goodman, C. D. Hodgson, and W. D. Neumann, "Computing arithmetic invariants of 3-manifolds", *Experiment. Math.* **9**:1 (2000), 127–152.

[Coxeter 1961] H. S. M. Coxeter, *Introduction to geometry*, John Wiley & Sons, New York, 1961.

[Culler 1986] M. Culler, "Lifting representations to covering groups", *Adv. in Math.* **59**:1 (1986), 64–70.

[Culler and Shalen 1983] M. Culler and P. B. Shalen, "Varieties of group representations and splittings of 3-manifolds", *Ann. of Math.* (2) **117**:1 (1983), 109–146.

[De-Spiller 1970] D. A. De-Spiller, "Equimorphisms, and quasiconformal mappings of the absolute", *Dokl. Akad. Nauk SSSR* **194** (1970), 1006–1009.

[Douady and Earle 1986] A. Douady and C. J. Earle, "Conformally natural extension of homeomorphisms of the circle", *Acta Math.* **157**:1-2 (1986), 23–48.

[Dumas 2004] D. Dumas, *Complex projective structures, grafting, and Teichmüller theory*, Ph.D. thesis, Harvard Univ., Cambridge (MA), 2004.

[Dunbar and Meyerhoff 1994] W. D. Dunbar and G. R. Meyerhoff, "Volumes of hyperbolic 3-orbifolds", *Indiana Univ. Math. J.* **43**:2 (1994), 611–637.

[Dunfield and Thurston 2003] N. M. Dunfield and W. P. Thurston, "The virtual Haken conjecture: experiments and examples", *Geom. Topol.* **7** (2003), 399–441.

[Duren 1983] P. L. Duren, *Univalent functions*, Grundlehren der Mathematischen Wissenschaften **259**, Springer, New York, 1983.

[Earle 1993] C. J. Earle, "The infinite Nielsen kernels of some bordered Riemann surfaces", *Michigan Math. J.* **40**:3 (1993), 445–458.

[Earle and Kra 1974] C. J. Earle and I. Kra, "On holomorphic mappings between Teichmüller spaces", pp. 107–124 in *Contributions to analysis (a collection of papers dedicated to Lipman Bers)*, Academic Press, New York, 1974.

[Earle et al. 1994] C. J. Earle, I. Kra, and S. L. Krushkal', "Holomorphic motions and Teichmüller spaces", *Trans. Amer. Math. Soc.* **343**:2 (1994), 927–948.

[Efremovič and Tihomirova 1964] V. A. Efremovič and E. S. Tihomirova, "Equimorphisms of hyperbolic spaces", *Izv. Akad. Nauk SSSR Ser. Mat.* **28** (1964), 1139–1144. In Russian.

[Epstein 1986] C. L. Epstein, "The hyperbolic Gauss map and quasiconformal reflections", *J. Reine Angew. Math.* **372** (1986), 96–135.

[Epstein and Gunn 1991] D. Epstein and C. Gunn, *Supplement to Not knot*, A K Peters, 1991.

[Epstein and Marden 1987] D. B. A. Epstein and A. Marden, "Convex hulls in hyperbolic space, a theorem of Sullivan, and measured pleated surfaces", pp. 113–253 in *Analytical and geometric aspects of hyperbolic space* (Warwick and Durham, 1984), edited by D. B. A. Epstein, London Math. Soc. Lecture Note Ser. **111**, Cambridge Univ. Press, Cambridge, 1987. Reprinted in [Canary et al. 2006].

[Epstein and Petronio 1994] D. B. A. Epstein and C. Petronio, "An exposition of Poincaré's polyhedron theorem", *Enseign. Math.* (2) **40**:1-2 (1994), 113–170.

[Epstein et al. 1992] D. B. A. Epstein, J. W. Cannon, D. F. Holt, S. V. F. Levy, M. S. Paterson, and W. P. Thurston, *Word processing in groups*, Jones and Bartlett Publishers, Boston, 1992.

[Epstein et al. 2004] D. B. A. Epstein, A. Marden, and V. Markovic, "Quasiconformal homeomorphisms and the convex hull boundary", *Ann. of Math.* (2) **159**:1 (2004), 305–336.

[Epstein et al. 2006] D. B. A. Epstein, A. Marden, and V. Markovic, "Complex earthquakes and deformations of the unit disk", *J. Diff. Geom* **73** (2006), 119–166.

[Evans 2004a] R. Evans, "Weakly type-preserving sequences and strong convergence", *Geom. Dedicata* **108** (2004), 71–92.

[Evans 2004b] R. A. Evans, "Tameness persists in weakly type-preserving strong limits", *Amer. J. Math.* **126**:4 (2004), 713–737.

[Evans 2006] R. A. Evans, "McMullen's conjecture for convex cores of hyperbolic 3-manifolds", preprint, 2006.

[Evans ≥ 2007] R. A. Evans, "Uniformly bounded radii of balls in convex cores of hyperbolic 3-manifolds", *Math. Annalen*. to appear.

[Farb 1998] B. Farb, "Relatively hyperbolic groups", *Geom. Funct. Anal.* **8**:5 (1998), 810–840.

[Farkas and Kra 1980] H. M. Farkas and I. Kra, *Riemann surfaces*, Graduate Texts in Mathematics **71**, Springer, New York, 1980.

[Fathi et al. 1979] A. Fathi, F. Laudenbach, and V. Poénaru, *Travaux de Thurston sur les surfaces*, Astérisque **66**, Société Mathématique de France, Paris, 1979.

[Fatou 1930] P. Fatou, "Fonctions automorphes", in *Théorie des fonctions algébriques et de leurs intégrales*, 2nd ed., vol. 2, Gauthier-Villars, Paris, 1930.

[Feighn and McCullough 1987] M. Feighn and D. McCullough, "Finiteness conditions for 3-manifolds with boundary", *Amer. J. Math.* **109**:6 (1987), 1155–1169.

[Feighn and Mess 1991] M. Feighn and G. Mess, "Conjugacy classes of finite subgroups of Kleinian groups", *Amer. J. Math.* **113**:1 (1991), 179–188.

[Fenchel 1989] W. Fenchel, *Elementary geometry in hyperbolic space*, Studies in Mathematics **11**, de Gruyter, Berlin, 1989.

[Floyd et al. 2002] W. Floyd, B. Weber, and J. Weeks, "The Achilles' heel of O(3, 1)?", *Experiment. Math.* **11**:1 (2002), 91–97.

[Ford 1929] L. R. Ford, *Automorphic Functions*, McGraw-Hill, New York, 1929. Reprinted by Chelsea, New York, 1951.

[Fox and Artin 1948] R. H. Fox and E. Artin, "Some wild cells and spheres in three-dimensional space", *Ann. of Math.* (2) **49** (1948), 979–990.

[Frigerio et al. 2004] R. Frigerio, B. Martelli, and C. Petronio, "Small hyperbolic 3-manifolds with geodesic boundary", *Experiment. Math.* **13**:2 (2004), 171–184.

[Fujii 1990] M. Fujii, "Hyperbolic 3-manifolds with totally geodesic boundary which are decomposed into hyperbolic truncated tetrahedra", *Tokyo J. Math.* **13**:2 (1990), 353–373.

[Gabai 1992] D. Gabai, "Convergence groups are Fuchsian groups", *Ann. of Math.* (2) **136**:3 (1992), 447–510.

[Gabai et al. 2001] D. Gabai, G. R. Meyerhoff, and P. Milley, "Volumes of tubes in hyperbolic 3-manifolds", *J. Differential Geom.* **57**:1 (2001), 23–46.

[Gabai et al. 2003] D. Gabai, G. R. Meyerhoff, and N. Thurston, "Homotopy hyperbolic 3-manifolds are hyperbolic", *Ann. of Math.* (2) **157**:2 (2003), 335–431.

[Gallo et al. 2000] D. Gallo, M. Kapovich, and A. Marden, "The monodromy groups of Schwarzian equations on closed Riemann surfaces", *Ann. of Math.* (2) **151**:2 (2000), 625–704.

[Gehring 1962] F. W. Gehring, "Rings and quasiconformal mappings in space", *Trans. Amer. Math. Soc.* **103** (1962), 353–393.

[Gehring and Martin 1994] F. W. Gehring and G. J. Martin, "Commutators, collars and the geometry of Möbius groups", *J. Anal. Math.* **63** (1994), 175–219.

[Gilman 1995] J. Gilman, *Two-generator discrete subgroups of* PSL(2, **R**), Mem. Amer. Math. Soc. **561**, Amer. Math. Soc., Providence, 1995.

[Gilman and Waterman 2003] J. Gilman and P. Waterman, "Parameters for two-parabolic Schottky groups", preprint, 2003.

[Goldman 1987] W. M. Goldman, "Projective structures with Fuchsian holonomy", *J. Differential Geom.* **25**:3 (1987), 297–326.

[Goldman 2003] W. M. Goldman, "The modular group action on real SL(2)-characters of a one-holed torus", *Geom. Topol.* **7** (2003), 443–486.

[Goodman n.d.] O. Goodman, "Snap", n.d. Available at http://www.ms.unimelb.edu.au/~snap. software to computing arithmetic invariants of hyperbolic 3-manifolds.

[Gordon and Luecke 1989] C. M. Gordon and J. Luecke, "Knots are determined by their complements", *J. Amer. Math. Soc.* **2**:2 (1989), 371–415.

[Gray 1986] J. Gray, *Linear differential equations and group theory from Riemann to Poincaré*, Birkhäuser, Boston, 1986. 2nd edition, 2000.

[Gray 2002] J. Gray, 2002. Unpublished manuscript.

[Greenberg 1960] L. Greenberg, "Discrete groups of motions", *Canad. J. Math.* **12** (1960), 415–426.

[Greenberg 1962] L. Greenberg, "Discrete subgroups of the Lorentz group", *Math. Scand.* **10** (1962), 85–107.

[Greenberg 1966] L. Greenberg, "Fundamental polyhedra for kleinian groups", *Ann. of Math.* (2) **84** (1966), 433–441.

[Greenberg 1974] L. Greenberg, "Commensurable groups of Moebius transformations", pp. 227–237 in *Discontinuous groups and Riemann surfaces* (College Park, MD, 1973), edited by L. Greenberg, Ann. of Math. Studies **79**, Princeton Univ. Press, Princeton, NJ, 1974.

[Greenberg 1977] L. Greenberg, "Finiteness theorems for Fuchsian and Kleinian groups", pp. 199–257 in *Discrete groups and automorphic functions* (Cambridge, 1975), edited by W. Harvey, Academic Press, London, 1977.

[Greenberg 1981] L. Greenberg, "Homomorphisms of triangle groups into PSL(2, **C**)", pp. 167–181 in *Riemann surfaces and related topics* (Stony Brook, NY, 1978), edited by I. Kra and B. Maskit, Ann. of Math. Stud. **97**, Princeton Univ. Press, Princeton, NJ, 1981.

[Gromov 1981a] M. Gromov, "Groups of polynomial growth and expanding maps", *Inst. Hautes Études Sci. Publ. Math.* no. 53 (1981), 53–73.

[Gromov 1981b] M. Gromov, "Hyperbolic manifolds (according to Thurston and Jørgensen)", pp. 40–53 in *Bourbaki Seminar, 1979/80*, Lecture Notes in Math. **842**, Springer, Berlin, 1981.

[Gromov 1987] M. Gromov, "Hyperbolic groups", pp. 75–263 in *Essays in group theory*, Math. Sci. Res. Inst. Publ. **8**, Springer, New York, 1987.

[Gromov and Thurston 1987] M. Gromov and W. Thurston, "Pinching constants for hyperbolic manifolds", *Invent. Math.* **89**:1 (1987), 1–12.

[Gunn and Maxwell 1991] C. Gunn and D. Maxwell, *Not knot*, A K Peters, 1991. Video.

[Harvey 1977] W. J. Harvey, "Spaces of discrete groups", pp. 295–348 in *Discrete groups and automorphic functions* (Cambridge, 1975), edited by W. Harvey, Academic Press, London, 1977.

[Harvey 1981] W. J. Harvey, "Boundary structure of the modular group", pp. 245–251 in *Riemann surfaces and related topics* (Stony Brook, NY, 1978), edited by I. Kra and B. Maskit, Ann. of Math. Stud. **97**, Princeton Univ. Press, Princeton, NJ, 1981.

[Hejhal 1975] D. A. Hejhal, "Monodromy groups and linearly polymorphic functions", *Acta Math.* **135**:1 (1975), 1–55.

[Hempel 1976] J. Hempel, 3-*Manifolds*, Princeton Univ. Press, Princeton, NJ, 1976.

[Hempel 1987] J. Hempel, "Residual finiteness for 3-manifolds", pp. 379–396 in *Combinatorial group theory and topology (Alta, Utah, 1984)*, Ann. of Math. Stud. **111**, Princeton Univ. Press, Princeton, NJ, 1987.

[Hildebrand and Weeks 1989] M. Hildebrand and J. Weeks, "A computer generated census of cusped hyperbolic 3-manifolds", pp. 53–59 in *Computers and mathematics (Cambridge, MA, 1989)*, Springer, New York, 1989.

[Hilden et al. 1985] H. M. Hilden, M. T. Lozano, and J. M. Montesinos, "On knots that are universal", *Topology* **24**:4 (1985), 499–504.

[Hocking and Young 1961] J. G. Hocking and G. S. Young, *Topology*, Addison-Wesley, Reading (MA), 1961.

[Hodgson and Kerckhoff 1998] C. D. Hodgson and S. P. Kerckhoff, "Rigidity of hyperbolic cone-manifolds and hyperbolic Dehn surgery", *J. Differential Geom.* **48**:1 (1998), 1–59.

[Hodgson and Weeks 1994] C. D. Hodgson and J. R. Weeks, "Symmetries, isometries and length spectra of closed hyperbolic three-manifolds", *Experiment. Math.* **3**:4 (1994), 261–274.

[Holt 2003] J. Holt, "Multiple bumping of components of deformation spaces of hyperbolic 3-manifolds", *Amer. J. Math.* **125**:4 (2003), 691–736.

[Imayoshi and Taniguchi 1992] Y. Imayoshi and M. Taniguchi, *An introduction to Teichmüller spaces*, Springer, Tokyo, 1992.

[Ito 2000a] K. Ito, "Exotic projective structures and quasi-Fuchsian space", *Duke Math. J.* **105**:2 (2000), 185–209.

[Ito 2000b] K. Ito, "Exotic projective structures and quasi-Fuchsian space", *Duke Math. J.* **105**:2 (2000), 185–209.

[Ivanov 1992] N. V. Ivanov, *Subgroups of Teichmüller modular groups*, Translations of Mathematical Monographs **115**, American Mathematical Society, Providence, RI, 1992.

[Jaco 1980] W. Jaco, *Lectures on three-manifold topology*, CBMS Regional Conference Series in Mathematics **43**, Amer. Math. Soc., Providence, 1980.

[Johannson 1979] K. Johannson, *Homotopy equivalences of* 3*-manifolds with boundaries*, Lecture Notes in Math. **761**, Springer, Berlin, 1979.

[Jones and Singerman 1978] G. A. Jones and D. Singerman, "Theory of maps on orientable surfaces", *Proc. London Math. Soc.* (3) **37**:2 (1978), 273–307.

[Jones and Singerman 1996] G. Jones and D. Singerman, "Belyĭ functions, hypermaps and Galois groups", *Bull. London Math. Soc.* **28**:6 (1996), 561–590.

[Jørgensen 1973] T. Jørgensen, "On cyclic groups of Möbius transformations", *Math. Scand.* **33** (1973), 250–260 (1974).

[Jørgensen 1974a] T. Jørgensen, "On reopening of cusps", 1974.

[Jørgensen 1974b] T. Jørgensen, "Some remarks on Kleinian groups", *Math. Scand.* **34** (1974), 101–108.

[Jørgensen 1976] T. Jørgensen, "On discrete groups of Möbius transformations", *Amer. J. Math.* **98**:3 (1976), 739–749.

[Jørgensen 1977a] T. Jørgensen, "Compact 3-manifolds of constant negative curvature fibering over the circle", *Ann. Math.* (2) **106**:1 (1977), 61–72.

[Jørgensen 1977b] T. Jørgensen, "A note on subgroups of $SL(2, \mathbf{C})$", *Quart. J. Math. Oxford Ser.* (2) **28**:110 (1977), 209–211.

[Jørgensen 1978] T. Jørgensen, "Closed geodesics on Riemann surfaces", *Proc. Amer. Math. Soc.* **72**:1 (1978), 140–142.

[Jørgensen 1981] T. Jørgensen, "Commutators in SL(2, **C**)", pp. 301–303 in *Riemann surfaces and related topics* (Stony Brook, NY, 1978), edited by I. Kra and B. Maskit, Ann. of Math. Stud. **97**, Princeton Univ. Press, Princeton, NJ, 1981.

[Jørgensen 2000] T. Jørgensen, "Composition and length of hyperbolic motions", pp. 211–220 in *In the tradition of Ahlfors and Bers* (Stony Brook, NY, 1998), Contemp. Math. **256**, Amer. Math. Soc., Providence, RI, 2000.

[Jørgensen 2003] T. Jørgensen, "On pairs of once-punctured tori", pp. 183–207 in *Kleinian groups and hyperbolic 3-manifolds (Warwick, 2001)*, London Math. Soc. Lecture Note Ser. **299**, Cambridge Univ. Press, Cambridge, 2003.

[Jørgensen and Kiikka 1975] T. Jørgensen and M. Kiikka, "Some extreme discrete groups", *Ann. Acad. Sci. Fenn. Ser. A I Math.* **1**:2 (1975), 245–248.

[Jørgensen and Klein 1982] T. Jørgensen and P. Klein, "Algebraic convergence of finitely generated Kleinian groups", *Quart. J. Math. Oxford Ser.* (2) **33**:131 (1982), 325–332.

[Jørgensen and Marden 1979] T. Jørgensen and A. Marden, "Two doubly degenerate groups", *Quart. J. Math. Oxford Ser.* (2) **30**:118 (1979), 143–156.

[Jørgensen and Marden 1990] T. Jørgensen and A. Marden, "Algebraic and geometric convergence of Kleinian groups", *Math. Scand.* **66**:1 (1990), 47–72.

[Jørgensen et al. 1992] T. Jørgensen, A. Lascurain, and T. Pignataro, "Translation extensions of the classical modular group", *Complex Variables Theory Appl.* **19**:4 (1992), 205–209.

[Kamishima and Tan 1992] Y. Kamishima and S. P. Tan, "Deformation spaces on geometric structures", pp. 263–299 in *Aspects of low-dimensional manifolds*, Adv. Stud. Pure Math. **20**, Kinokuniya, Tokyo, 1992.

[Kapovich 2001] M. Kapovich, *Hyperbolic manifolds and discrete groups*, Progress in Mathematics **183**, Birkhäuser, Boston, 2001.

[Kazhdan and Margulis 1968] D. A. Kazhdan and G. A. Margulis, “A proof of Selberg’s conjecture”, *Mat. Sb.* (*N.S.*) **75**:(117) (1968), 163–168. In Russian; translated in *Math. USSR Sbornik* **4** (1968), 147-152.

[Kellerhals 1998] R. Kellerhals, “Volumes of cusped hyperbolic manifolds”, *Topology* **37**:4 (1998), 719–734.

[Kellerhals and Zehrt 2001] R. Kellerhals and T. Zehrt, “The Gauss-Bonnet formula for hyperbolic manifolds of finite volume”, *Geom. Dedicata* **84**:1-3 (2001), 49–62.

[Kerckhoff 1980] S. P. Kerckhoff, “The asymptotic geometry of Teichmüller space”, *Topology* **19**:1 (1980), 23–41.

[Kerckhoff 1983] S. P. Kerckhoff, “The Nielsen realization problem”, *Ann. of Math.* (2) **117**:2 (1983), 235–265.

[Kerckhoff 1985] S. P. Kerckhoff, “Earthquakes are analytic”, *Comment. Math. Helv.* **60**:1 (1985), 17–30.

[Kerckhoff 1990] S. P. Kerckhoff, “The measure of the limit set of the handlebody group”, *Topology* **29**:1 (1990), 27–40.

[Kerckhoff and Thurston 1990] S. P. Kerckhoff and W. P. Thurston, “Noncontinuity of the action of the modular group at Bers’ boundary of Teichmüller space”, *Invent. Math.* **100**:1 (1990), 25–47.

[Klarreich 1999] E. Klarreich, “Semiconjugacies between Kleinian group actions on the Riemann sphere”, *Amer. J. Math.* **121**:5 (1999), 1031–1078.

[Klein 1871] F. Klein, “Über die sogenannte nicht-euklidische Geometrie”, *Math. Ann.* **4** (1871), 573–625. Page number cited in text is for the English translation, which appears in *Sources of hyperbolic geometry*, translated and edited by John Stillwell, Amer. Math. Soc., Providence, 1996.

[Kleineidam and Souto 2002] G. Kleineidam and J. Souto, “Algebraic convergence of function groups”, *Comment. Math. Helv.* **77**:2 (2002), 244–269.

[Kleineidam and Souto 2003] G. Kleineidam and J. Souto, “Ending laminations in the Masur domain”, pp. 105–129 in *Kleinian groups and hyperbolic 3-manifolds (Warwick, 2001)*, London Math. Soc. Lecture Note Ser. **299**, Cambridge Univ. Press, Cambridge, 2003.

[Kojima 1988] S. Kojima, “Isometry transformations of hyperbolic 3-manifolds”, *Topology Appl.* **29**:3 (1988), 297–307.

[Kojima 1996] S. Kojima, “Nonsingular parts of hyperbolic 3-cone-manifolds”, pp. 115–122 in *Topology and Teichmüller spaces* (Katinkulta, 1995), edited by S. Kojima et al., World Scientific, River Edge, NJ, 1996.

[Kojima 1998] S. Kojima, “Deformations of hyperbolic 3-cone-manifolds”, *J. Differential Geom.* **49**:3 (1998), 469–516.

[Kojima et al. 2006] S. Kojima, S. Mizushima, and S. P. Tan, “Circle packings on surfaces with projective structures: A survey”, in *Spaces of kleinian groups*, edited by Y. Minsky et al., London Math. Soc. Lecture Notes **329**, Camb. Univ. Press, 2006.

[Komori and Sugawa 2004] Y. Komori and T. Sugawa, “Bers embedding of the Teichmüller space of a once-punctured torus”, *Conform. Geom. Dyn.* **8** (2004), 115–142.

[Koundouros 2004] S. Koundouros, “Universal surgery bounds on hyperbolic 3-manifolds”, *Topology* **43**:3 (2004), 497–512.

[Kra 1972] I. Kra, “On spaces of Kleinian groups”, *Comment. Math. Helv.* **47** (1972), 53–69.

[Kulkarni and Shalen 1989] R. S. Kulkarni and P. B. Shalen, “On Ahlfors’ finiteness theorem”, *Adv. Math.* **76**:2 (1989), 155–169.

[Lamping et al. 1995] J. Lamping, R. Rao, and P. Pirolli, “A focus + content technique based on hyperbolic geometry for viewing large hierarchies”, pp. 401–408 in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, ACM, New York, 1995.

[Lecuire 2003] C. Lecuire, “Une caractérisation des laminations géodésiques mesurées de plissage des variétés hyperboliques et ses conséquences”, pp. 103–115 in *Séminaire de Théorie Spectrale et Géométrie, vol. 21. Année 2002–2003*, Sémin. Théor. Spectr. Géom. **21**, Univ. Grenoble I, Saint, 2003.

[Lecuire 2004a] C. Lecuire, "Bending map and strong convergence", preprint, University of Warwick, 2004. Available at http://www.maths.warwick.ac.uk/~clecuire/.

[Lecuire 2004b] C. Lecuire, "Continuity of the bending map", preprint, University of Warwick, 2004. Available at arXiv.org/abs/math.DG/0411412.

[Lecuire 2004c] C. Lecuire, "An extension of Masur domain", preprint, University of Warwick, 2004.

[Lehner 1964] J. Lehner, *Discontinuous groups and automorphic functions*, Mathematical Surveys **8**, Amer. Math. Soc., Providence, 1964.

[Lehto 1987] O. Lehto, *Univalent functions and Teichmüller spaces*, Graduate Texts in Mathematics **109**, Springer, New York, 1987.

[Leung and Wan 2001] N. C. Leung and T. Y. H. Wan, "Harmonic maps and the topology of conformally compact Einstein manifolds", *Math. Res. Lett.* **8**:5-6 (2001), 801–812.

[Lickorish 1997] W. B. R. Lickorish, *An introduction to knot theory*, Graduate Texts in Mathematics **175**, Springer, New York, 1997.

[Maclachlan and Reid 2003] C. Maclachlan and A. W. Reid, *The arithmetic of hyperbolic 3-manifolds*, Graduate Texts in Mathematics **219**, Springer, New York, 2003.

[Magnus 1974] W. Magnus, *Noneuclidean tesselations and their groups*, Pure and Applied Mathematics **61**, Academic Press, New York, 1974.

[Magnus 1980] W. Magnus, "Rings of Fricke characters and automorphism groups of free groups", *Math. Z.* **170**:1 (1980), 91–103.

[Marden 1967] A. Marden, "On finitely generated Fuchsian groups", *Comment. Math. Helv.* **42** (1967), 81–85.

[Marden 1969] A. Marden, "On homotopic mappings of Riemann surfaces", *Ann. of Math.* (2) **90** (1969), 1–8.

[Marden 1971] A. Marden, "An inequality for Kleinian groups", pp. 295–296 in *Advances in the Theory of Riemann Surfaces (Proc. Conf., Stony Brook, N.Y., 1969)*, Ann. of Math. Studies, No. 66, Princeton Univ. Press, Princeton, NJ, 1971.

[Marden 1974a] A. Marden, "The geometry of finitely generated kleinian groups", *Ann. of Math.* (2) **99** (1974), 383–462.

[Marden 1974b] A. Marden, "Kleinian groups and 3-dimensional topology: a survey", pp. 108–121 in *A crash course on Kleinian groups* (San Francisco, CA, 1974), edited by L. Bers and I. Kra, Lecture Notes in Math. **400**, Springer, Berlin, 1974.

[Marden 1974c] A. Marden, "Schottky groups and circles", pp. 273–278 in *Contributions to analysis (a collection of papers dedicated to Lipman Bers)*, Academic Press, New York, 1974.

[Marden 1974d] A. Marden, "Universal properties of Fuchsian groups in the Poincaré metric", pp. 315–339 in *Discontinuous groups and Riemann surfaces* (College Park, MD, 1973), edited by L. Greenberg, Ann. of Math. Studies **79**, Princeton Univ. Press, Princeton, NJ, 1974.

[Marden 1977] A. Marden, "Geometrically finite Kleinian groups and their deformation spaces", pp. 259–293 in *Discrete groups and automorphic functions* (Cambridge, 1975), edited by W. Harvey, Academic Press, London, 1977.

[Marden 1980] A. Marden, "Geometric relations between homeomorphic Riemann surfaces", *Bull. Amer. Math. Soc.* (*N.S.*) **3**:3 (1980), 1001–1017.

[Marden 2006] A. Marden, "A proof of the Ahlfors Finiteness Theorem", in *Spaces of kleinian groups*, edited by Y. Minsky et al., London Math. Soc. Lecture Notes **329**, Camb. Univ. Press, 2006.

[Marden $\geq$ 2007] A. Marden, "Handbook on Teichmueller spaces", European Math. Soc. to appear.

[Marden and Strebel 1984] A. Marden and K. Strebel, "The heights theorem for quadratic differentials on Riemann surfaces", *Acta Math.* **153**:3-4 (1984), 153–211.

[Marden and Strebel 1985] A. Marden and K. Strebel, "On the ends of trajectories", pp. 195–204 in *Differential geometry and complex analysis*, Springer, Berlin, 1985.

[Marden and Strebel 1986] A. Marden and K. Strebel, “Pseudo-Anosov Teichmueller mappings”, *J. Analyse Math.* **46** (1986), 194–220.

[Marden and Strebel 1993] A. Marden and K. Strebel, “A characterization of Teichmüller differentials”, *J. Differential Geom.* **37**:1 (1993), 1–29.

[Markovic 2005] V. Markovic, “Realization of the mapping class group by homeomorphisms”, preprint, University of Warwick, 2005. Available at http://www.maths.warwick.ac.uk/~markovic.

[Markovic and Sarić 2004] V. Markovic and D. Sarić, “Teichmüller mapping class group of the universal hyperbolic solenoid”, preprint, 2004.

[Marshall and Martin ≥ 2007] T. H. Marshall and G. J. Martin, to appear.

[Maskit 1967] B. Maskit, “A characterization of Schottky groups”, *J. Analyse Math.* **19** (1967), 227–230.

[Maskit 1968] B. Maskit, “The conformal group of a plane domain”, *Amer. J. Math.* **90** (1968), 718–722.

[Maskit 1970] B. Maskit, “On boundaries of Teichmüller spaces and on Kleinian groups, II”, *Ann. of Math.* (2) **91** (1970), 607–639.

[Maskit 1971] B. Maskit, “Self-maps on Kleinian groups”, *Amer. J. Math.* **93** (1971), 840–856.

[Maskit 1974] B. Maskit, “Intersections of component subgroups of Kleinian groups”, pp. 349–367 in *Discontinuous groups and Riemann surfaces* (College Park, MD, 1973), edited by L. Greenberg, Ann. of Math. Studies **79**, Princeton Univ. Press, Princeton, NJ, 1974.

[Maskit 1983] B. Maskit, “Parabolic elements in Kleinian groups”, *Ann. of Math.* (2) **117**:3 (1983), 659–668.

[Maskit 1988] B. Maskit, *Kleinian groups*, Grundlehren der Mathematischen Wissenschaften **287**, Springer, Berlin, 1988.

[Masur 1975] H. Masur, “On a class of geodesics in Teichmüller space”, *Ann. of Math.* (2) **102**:2 (1975), 205–221.

[Masur 1982] H. Masur, “Interval exchange transformations and measured foliations”, *Ann. of Math.* (2) **115**:1 (1982), 169–200.

[Masur 1986] H. Masur, “Measured foliations and handlebodies”, *Ergodic Theory Dynam. Systems* **6**:1 (1986), 99–116.

[Masur 1992] H. Masur, “Hausdorff dimension of the set of nonergodic foliations of a quadratic differential”, *Duke Math. J.* **66**:3 (1992), 387–442.

[Masur and Minsky 1999] H. A. Masur and Y. N. Minsky, “Geometry of the complex of curves, I: Hyperbolicity”, *Invent. Math.* **138**:1 (1999), 103–149.

[Masur and Minsky 2000] H. A. Masur and Y. N. Minsky, “Geometry of the complex of curves, II: Hierarchical structure”, *Geom. Funct. Anal.* **10**:4 (2000), 902–974.

[Masur and Tabachnikov 2002] H. Masur and S. Tabachnikov, “Rational billiards and flat structures”, pp. 1015–1089 in *Handbook of dynamical systems*, vol. 1A, North-Holland, Amsterdam, 2002.

[Matsuzaki and Taniguchi 1998] K. Matsuzaki and M. Taniguchi, *Hyperbolic manifolds and Kleinian groups*, Oxford Mathematical Monographs, The Clarendon Press Oxford Univ. Press, New York, 1998.

[McCullough 1986] D. McCullough, “Compact submanifolds of 3-manifolds with boundary”, *Quart. J. Math. Oxford Ser.* (2) **37**:147 (1986), 299–307.

[McCullough and Miller 1986] D. McCullough and A. Miller, *Homeomorphisms of 3-manifolds with compressible boundary*, Mem. Amer. Math. Soc. **344**, Amer. Math. Soc., Providence, 1986.

[McCullough et al. 1985] D. McCullough, A. Miller, and G. A. Swarup, “Uniqueness of cores of noncompact 3-manifolds”, *J. London Math. Soc.* (2) **32**:3 (1985), 548–556.

[McMullen 1990] C. McMullen, “Iteration on Teichmüller space”, *Invent. Math.* **99**:2 (1990), 425–454.

[McMullen 1991] C. McMullen, “Cusps are dense”, *Ann. of Math.* (2) **133**:1 (1991), 217–247.

[McMullen 1996] C. T. McMullen, *Renormalization and 3-manifolds which fiber over the circle*, Annals of Mathematics Studies **142**, Princeton Univ. Press, Princeton, NJ, 1996.

[McMullen 1998] C. T. McMullen, "Complex earthquakes and Teichmüller theory", *J. Amer. Math. Soc.* **11**:2 (1998), 283–320.

[McMullen 1999] C. T. McMullen, "Hausdorff dimension and conformal dynamics, I: Strong convergence of Kleinian groups", *J. Differential Geom.* **51**:3 (1999), 471–515.

[McMullen 2001] C. T. McMullen, "Local connectivity, Kleinian groups and geodesics on the blowup of the torus", *Invent. Math.* **146**:1 (2001), 35–91.

[McShane 1998] G. McShane, "Simple geodesics and a series constant over Teichmuller space", *Invent. Math.* **132**:3 (1998), 607–632.

[Meeks and Yau 1981] W. H. Meeks, III and S. T. Yau, "The equivariant Dehn's lemma and loop theorem", *Comment. Math. Helv.* **56**:2 (1981), 225–239.

[Meyerhoff 1987] R. Meyerhoff, "A lower bound for the volume of hyperbolic 3-manifolds", *Canad. J. Math.* **39**:5 (1987), 1038–1056.

[Milnor 1994] J. Milnor, *Collected papers*, vol. 1, Publish or Perish, Houston, TX, 1994.

[Milnor 2003] J. Milnor, "Towards the Poincaré conjecture and the classification of 3-manifolds", *Notices Amer. Math. Soc.* **50**:10 (2003), 1226–1233.

[Minsky 1992] Y. N. Minsky, "Harmonic maps, length, and energy in Teichmüller space", *J. Differential Geom.* **35**:1 (1992), 151–217.

[Minsky 1993] Y. N. Minsky, "Teichmüller geodesics and ends of hyperbolic 3-manifolds", *Topology* **32**:3 (1993), 625–647.

[Minsky 1994a] Y. N. Minsky, "On rigidity, limit sets, and end invariants of hyperbolic 3-manifolds", *J. Amer. Math. Soc.* **7**:3 (1994), 539–588.

[Minsky 1994b] Y. N. Minsky, "On Thurston's ending lamination conjecture", pp. 109–122 in *Low-dimensional topology* (Knoxville, TN, 1992), edited by K. Johannson, Conf. Proc. Lecture Notes Geom. Topology **3**, Internat. Press, Cambridge, MA, 1994.

[Minsky 1999] Y. N. Minsky, "The classification of punctured-torus groups", *Ann. of Math.* (2) **149**:2 (1999), 559–626.

[Minsky 2000] Y. Minsky, "Kleinian groups and the complex of curves", *Geom. Top.* **4** (2000), 117–148.

[Minsky 2001] Y. N. Minsky, "Bounded geometry for Kleinian groups", *Invent. Math.* **146**:1 (2001), 143–192.

[Minsky 2003a] Y. Minsky, "Classification of Kleinian surface groups, I: Models and bounds", preprint, 2003. Available at www.arXiv.org/abs/math.GT/0302208.

[Minsky 2003b] Y. N. Minsky, "Combinatorial and geometrical aspects of hyperbolic 3-manifolds", pp. 3–40 in *Kleinian groups and hyperbolic 3-manifolds (Warwick, 2001)*, London Math. Soc. Lecture Note Ser. **299**, Cambridge Univ. Press, Cambridge, 2003.

[Mirzakhani 2004] M. Mirzakhani, "Growth of the number of simple closed geodesics on hyperbolic surfaces", *Ann. Math.* (2004). To appear.

[Mitra 1998a] M. Mitra, "Cannon-Thurston maps for hyperbolic group extensions", *Topology* **37**:3 (1998), 527–538.

[Mitra 1998b] M. Mitra, "Cannon-Thurston maps for trees of hyperbolic metric spaces", *J. Differential Geom.* **48**:1 (1998), 135–164.

[Miyachi 2003] H. Miyachi, "Cusps in complex boundaries of one-dimensional Teichmüller space", *Conform. Geom. Dyn.* **7** (2003), 103–151.

[Montesinos 1987] J. M. Montesinos, *Classical tessellations and three-manifolds*, Universitext, Springer, Berlin, 1987.

[Morgan 1984] J. W. Morgan, "On Thurston's uniformization theorem for three-dimensional manifolds", pp. 37–125 in *The Smith conjecture (New York, 1979)*, Pure Appl. Math. **112**, Academic Press, Orlando, FL, 1984.

[Morgan 2005] J. W. Morgan, "Recent progress on the Poincaré conjecture and the classification of 3-manifolds", *Bull. Amer. Math. Soc.* (*N.S.*) **42**:1 (2005), 57–78.

[Morgan and Shalen 1984] J. W. Morgan and P. B. Shalen, "Valuations, trees, and degenerations of hyperbolic structures, I", *Ann. of Math.* (2) **120**:3 (1984), 401–476.

[Morgan and Shalen 1988a] J. W. Morgan and P. B. Shalen, "Degenerations of hyperbolic structures, II: Measured laminations in 3-manifolds", *Ann. of Math.* (2) **127**:2 (1988), 403–456.

[Morgan and Shalen 1988b] J. W. Morgan and P. B. Shalen, "Degenerations of hyperbolic structures, III: Actions of 3-manifold groups on trees and Thurston's compactness theorem", *Ann. of Math.* (2) **127**:3 (1988), 457–519.

[Mosher $\geq$ 2007] L. Mosher, "Geometric survey of subgroups of mapping class groups", European Math. Soc. to appear.

[Mostow 1973] G. D. Mostow, *Strong rigidity of locally symmetric spaces*, Princeton Univ. Press, Princeton, NJ, 1973.

[Mulase and Penkava 1998] M. Mulase and M. Penkava, "Ribbon graphs, quadratic differentials on Riemann surfaces, and algebraic curves defined over $\overline{\mathbf{Q}}$", *Asian J. Math.* **2**:4 (1998), 875–919.

[Mumford 1971] D. Mumford, "A remark on Mahler's compactness theorem", *Proc. Amer. Math. Soc.* **28** (1971), 289–294.

[Mumford et al. 2002] D. Mumford, C. Series, and D. Wright, *Indra's pearls*, Cambridge Univ. Press, New York, 2002.

[Munkres 1960] J. Munkres, "Obstructions to the smoothing of piecewise-differentiable homeomorphisms", *Ann. of Math.* (2) **72** (1960), 521–554.

[Munzner 1997] T. Munzner, "H3: Laying out large directed graphs in hyperbolic space", pp. 2–10 in *Proc. 1997 IEEE Symposium on Information Visualization*, 1997.

[Myers 1982] R. Myers, "Simple knots in compact, orientable 3-manifolds", *Trans. Amer. Math. Soc.* **273**:1 (1982), 75–91.

[Myers 2005] R. Myers, "End reductions, fundamental groups, and covering spaces of irreducible open 3-manifolds", *Geom. Topol.* **9** (2005), 971–990.

[Neumann 1999] W. D. Neumann, "Notes on geometry and 3-manifolds", pp. 191–267 in *Low dimensional topology* (Eger, 1996, and Budapest, 1998), edited by J. Károly Böröczky et al., Bolyai Soc. Math. Stud. **8**, János Bolyai Math. Soc., Budapest, 1999.

[Neumann and Reid 1992] W. D. Neumann and A. W. Reid, "Arithmetic of hyperbolic manifolds", pp. 273–310 in *Topology '90* (Columbus, OH, 1990), edited by B. Apanasov et al., de Gruyter, Berlin, 1992.

[Nicholls 1989] P. J. Nicholls, *The ergodic theory of discrete groups*, London Mathematical Society Lecture Note Series **143**, Cambridge Univ. Press, Cambridge, 1989.

[Ohshika 1990] K. Ohshika, "Ending laminations and boundaries for deformation spaces of Kleinian groups", *J. London Math. Soc.* (2) **42**:1 (1990), 111–121.

[Ohshika 1998a] K. Ohshika, "Geometrically finite Kleinian groups and parabolic elements", *Proc. Edinburgh Math. Soc.* (2) **41**:1 (1998), 141–159.

[Ohshika 1998b] K. Ohshika, "Rigidity and topological conjugates of topologically tame Kleinian groups", *Trans. Amer. Math. Soc.* **350**:10 (1998), 3989–4022.

[Ohshika 2002] K. Ohshika, *Discrete groups*, Translations of Mathematical Monographs **207**, Amer. Math. Soc., Providence, RI, 2002.

[Ohshika 2003] K. Ohshika, "Constructing geometrically infinite groups on boundaries of deformation spaces", preprint, 2003.

[Ohshika 2004] K. Ohshika, "Realising end invariants by limits of convex cocompact groups", preprint, 2004.

[Ohshika 2005] K. Ohshika, *Kleinian groups which are limits of geometrically finite groups*, Mem. Amer. Math. Soc. **834**, Amer. Math. Soc., Providence, 2005.

[Ol'shanskiĭ 1992] A. Y. Ol'shanskiĭ, "Almost every group is hyperbolic", *Internat. J. Algebra Comput.* **2**:1 (1992), 1–17.

[Otal 1995] J.-P. Otal, "Sur le nouage des géodésiques dans les variétés hyperboliques", *C. R. Acad. Sci. Paris Sér. I Math.* **320**:7 (1995), 847–852.

[Otal 1996] J.-P. Otal, *Le théorème d'hyperbolisation pour les variétés fibrées de dimension 3*, Astérisque **235**, 1996. English translation, *The hyperbolization theorem for fibered* *manifolds of dimension 3*, SMF/AMS Texts and Monographs **7**, Amer. Math. Soc., Providence, RI, 2001.

[Otal 1998] J.-P. Otal, *Hyperbolization of* 3*-manifolds*, vol. III, edited by C. C. Hsiung and S.-T. Yau, International Press, Boston, 1998.

[Parker and Series 1995] J. R. Parker and C. Series, "Bending formulae for convex hull boundaries", *J. Anal. Math.* **67** (1995), 165–198.

[Patterson 1987] S. J. Patterson, "Lectures on measures on limit sets of Kleinian groups", pp. 281–323 in *Analytical and geometric aspects of hyperbolic space* (Warwick and Durham, 1984), edited by D. B. A. Epstein, London Math. Soc. Lecture Note Ser. **111**, Cambridge Univ. Press, Cambridge, 1987.

[Penner and Harer 1992] R. C. Penner and J. L. Harer, *Combinatorics of train tracks*, Annals of Mathematics Studies **125**, Princeton Univ. Press, Princeton, NJ, 1992.

[Petronio and Porti 2000] C. Petronio and J. Porti, "Negatively oriented ideal triangulations and a proof of Thurston's hyperbolic Dehn filling theorem", *Expo. Math.* **18**:1 (2000), 1–35.

[Petronio and Weeks 2000] C. Petronio and J. R. Weeks, "Partially flat ideal triangulations of cusped hyperbolic 3-manifolds", *Osaka J. Math.* **37**:2 (2000), 453–466.

[Pommerenke 1984] C. Pommerenke, "On uniformly perfect sets and Fuchsian groups", *Analysis* **4**:3-4 (1984), 299–321.

[Pommerenke 1992] C. Pommerenke, *Boundary behaviour of conformal maps*, Grundlehren der Mathematischen Wissenschaften **299**, Springer, Berlin, 1992.

[Prasad 1973] G. Prasad, "Strong rigidity of **Q**-rank 1 lattices", *Invent. Math.* **21** (1973), 255–286.

[Przeworski 2003] A. Przeworski, "Tubes in hyperbolic 3-manifolds", *Topology Appl.* **128**:2-3 (2003), 103–122.

[Ratcliffe 1994] J. G. Ratcliffe, *Foundations of hyperbolic manifolds*, Graduate Texts in Mathematics **149**, Springer, New York, 1994.

[Ratcliffe and Tschantz 2000] J. G. Ratcliffe and S. T. Tschantz, "The volume spectrum of hyperbolic 4-manifolds", *Experiment. Math.* **9**:1 (2000), 101–125.

[Rees 1981] M. Rees, "An alternative approach to the ergodic theory of measured foliations on surfaces", *Ergodic Theory Dynamical Systems* **1**:4 (1981), 461–488 (1982).

[Reimann 1985] H. M. Reimann, "Invariant extension of quasiconformal deformations", *Ann. Acad. Sci. Fenn. Ser. A I Math.* **10** (1985), 477–492.

[Riley 1975] R. Riley, "A quadratic parabolic group", *Math. Proc. Cambridge Philos. Soc.* **77** (1975), 281–288.

[Riley 1983] R. Riley, "Applications of a computer implementation of Poincaré's theorem on fundamental polyhedra", *Math. Comp.* **40**:162 (1983), 607–632.

[Rivin 1996] I. Rivin, "A characterization of ideal polyhedra in hyperbolic 3-space", *Ann. of Math.* (2) **143**:1 (1996), 51–70.

[Rodin and Sullivan 1987] B. Rodin and D. Sullivan, "The convergence of circle packings to the Riemann mapping", *J. Differential Geom.* **26**:2 (1987), 349–360.

[Rolfsen 1976] D. Rolfsen, *Knots and links*, Math. Lecture Series **7**, Publish and Perish, Berkeley, 1976. Second printing, 1990.

[Royden 1971] H. L. Royden, "Automorphisms and isometries of Teichmüller space", pp. 369–383 in *Advances in the theory of Riemann surfaces* (Stony Brook, NY, 1969), edited by L. Ahlfors et al., Ann. of Math. Studies **66**, Princeton Univ. Press, Princeton, NJ, 1971.

[Rüedy 1971] R. A. Rüedy, "Embeddings of open Riemann surfaces", *Comment. Math. Helv.* **46** (1971), 214–225.

[Scannell and Wolf 2002] K. P. Scannell and M. Wolf, "The grafting map of Teichmüller space", *J. Amer. Math. Soc.* **15**:4 (2002), 893–927.

[Schafer 1976] J. A. Schafer, "Representing homology classes on surfaces", *Canad. Math. Bull.* **19**:3 (1976), 373–374.

[Schneps 1994] L. Schneps (editor), *The Grothendieck theory of dessins d'enfants*, edited by L. Schneps, London Mathematical Society Lecture Note Series **200**, Cambridge Univ. Press, Cambridge, 1994.

[Schwartz 2006] R. E. Schwartz, "A conformal averaging process on the circle", *Geometriae Dedicata* **117** (2006), 19–46.

[Scott 1973a] G. P. Scott, "Compact submanifolds of 3-manifolds", *J. London Math. Soc.* (2) **7** (1973), 246–250.

[Scott 1973b] G. P. Scott, "Finitely generated 3-manifold groups are finitely presented", *J. London Math. Soc.* (2) **6** (1973), 437–440.

[Scott 1978] P. Scott, "Subgroups of surface groups are almost geometric", *J. London Math. Soc.* (2) **17**:3 (1978), 555–565.

[Scott 1980] P. Scott, "A new proof of the annulus and torus theorems", *Amer. J. Math.* **102**:2 (1980), 241–277.

[Scott 1983] P. Scott, "The geometries of 3-manifolds", *Bull. London Math. Soc.* **15**:5 (1983), 401–487.

[Scott and Tucker 1989] P. Scott and T. Tucker, "Some examples of exotic noncompact 3-manifolds", *Quart. J. Math. Oxford Ser.* (2) **40**:160 (1989), 481–499.

[Selberg 1960] A. Selberg, "On discontinuous groups in higher-dimensional symmetric spaces", pp. 147–164 in *Contributions to function theory (internat. Colloq. Function Theory, Bombay, 1960)*, Tata Institute of Fundamental Research, Bombay, 1960.

[Series 2004] C. Series, "Thurston's bending measure conjecture for once punctured torus groups", preprint, University of Warwick, 2004. Available at arXiv.org/abs/math.DG/0406056.

[Shiga and Tanigawa 1999] H. Shiga and H. Tanigawa, "Projective structures with discrete holonomy representations", *Trans. Amer. Math. Soc.* **351**:2 (1999), 813–823.

[Shimizu 1963] H. Shimizu, "On discontinuous groups operating on the product of the upper half planes", *Ann. of Math.* (2) **77** (1963), 33–71.

[Slodkowski 1991] Z. Slodkowski, "Holomorphic motions and polynomial hulls", *Proc. Amer. Math. Soc.* **111**:2 (1991), 347–355.

[Soma 2003] T. Soma, "Geometric limits of quasi-fuchsian groups", preprint, 2003.

[Soma 2005] T. Soma, "Existence of polygonal wrappings in hyperbolic 3-manifolds", preprint, 2005.

[Souto 2005] J. Souto, "A note on the tameness of hyperbolic 3-manifolds", *Topology* **44**:2 (2005), 459–474.

[Souto 2006] J. Souto, "Rank and topology of hyperbolic 3-manifolds, I", preprint, 2006.

[Springer 1957] G. Springer, *Introduction to Riemann surfaces*, Addison-Wesley, Reading, MA, 1957.

[Stephenson 1999] K. Stephenson, "The approximation of conformal structures via circle packing", pp. 551–582 in *Computational methods and function theory* (Nicosia, 1997), Ser. Approx. Decompos. **11**, World Scientific, River Edge, NJ, 1999.

[Stephenson 2003] K. Stephenson, "Circle packing: a mathematical tale", *Notices Amer. Math. Soc.* **50**:11 (2003), 1376–1388.

[Stephenson 2005] K. Stephenson, *Introduction to circle packing*, Cambridge Univ. Press, Cambridge, 2005.

[Stillwell 1996] J. Stillwell, *Sources of hyperbolic geometry*, History of Mathematics **10**, Amer. Math. Soc., Providence, RI, 1996.

[Storm 2002a] P. Storm, "The barycenter method on singular spaces", preprint, 2002. Available at www.arXiv.org/abs/math.GT/0301087.

[Storm 2002b] P. A. Storm, "Minimal volume Alexandrov spaces", *J. Differential Geom.* **61**:2 (2002), 195–225.

[Strebel 1984] K. Strebel, *Quadratic differentials*, vol. 5, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], Springer, Berlin, 1984.

[Struik 1950] D. J. Struik, *Lectures on classical differential geometry*, Addison-Wesley, Cambridge, MA, 1950.

[Sturm and Shinnar 1974] J. Sturm and M. Shinnar, "The maximal inscribed ball of a Fuchsian group", pp. 439–443 in *Discontinuous groups and Riemann surfaces* (College Park, MD, 1973), edited by L. Greenberg, Ann. of Math. Studies **79**, Princeton Univ. Press, Princeton, NJ, 1974.

[Sullivan 1981] D. Sullivan, "On the ergodic theory at infinity of an arbitrary discrete group of hyperbolic motions", pp. 465–496 in *Riemann surfaces and related topics* (Stony Brook, NY, 1978), edited by I. Kra and B. Maskit, Ann. of Math. Stud. **97**, Princeton Univ. Press, Princeton, NJ, 1981.

[Sullivan 1985] D. Sullivan, "Quasiconformal homeomorphisms and dynamics, II: Structural stability implies hyperbolicity for Kleinian groups", *Acta Math.* **155**:3-4 (1985), 243–260.

[Sullivan 1987] D. Sullivan, "Related aspects of positivity in Riemannian geometry", *J. Differential Geom.* **25**:3 (1987), 327–351.

[Swarup 1980] G. A. Swarup, "Two finiteness properties in 3-manifolds", *Bull. London Math. Soc.* **12**:4 (1980), 296–302.

[Tanigawa 1997] H. Tanigawa, "Grafting, harmonic maps and projective structures on surfaces", *J. Differential Geom.* **47**:3 (1997), 399–419.

[Thurston 1979] W. P. Thurston, "The geometry and topology of three-manifolds", lecture notes, Princeton University, 1979. Available at http://msri.org/publications/books/gt3m.

[Thurston 1980] W. P. Thurston, "Hyperbolic structures on 3-manifolds: overall logic", lecture notes, Bowdoin College, 1980.

[Thurston 1982a] W. Thurston, "Hyperbolic geometry and 3-manifolds", pp. 9–25 in *Low-dimensional topology (Bangor, 1979)*, London Math. Soc. Lecture Note Ser. **48**, Cambridge Univ. Press, Cambridge, 1982.

[Thurston 1982b] W. P. Thurston, "Three-dimensional manifolds, Kleinian groups and hyperbolic geometry", *Bull. Amer. Math. Soc.* (*N.S.*) **6**:3 (1982), 357–381.

[Thurston 1986a] W. P. Thurston, "Earthquakes in two-dimensional hyperbolic geometry", pp. 91–112 in *Analytical and geometric aspects of hyperbolic space* (Warwick and Durham, 1984), edited by D. B. A. Epstein, London Math. Soc. Lecture Note Ser. **112**, Cambridge Univ. Press, Cambridge, 1986. Reprinted in [Canary et al. 2006].

[Thurston 1986b] W. P. Thurston, "Hyperbolic structures on 3-manifolds, I: Deformation of acylindrical manifolds", *Ann. of Math.* (2) **124**:2 (1986), 203–246.

[Thurston 1986c] W. P. Thurston, "Hyperbolic structures on 3-manifolds, II: surface groups and 3-manifolds which fiber over the circle", preprint, 1986. Available at www.arXiv.org/abs/math.GT/9801045.

[Thurston 1986d] W. P. Thurston, "Zippers and univalent functions", pp. 185–197 in *The Bieberbach conjecture (West Lafayette, Ind., 1985)*, Math. Surveys Monogr. **21**, Amer. Math. Soc., Providence, RI, 1986.

[Thurston 1988] W. P. Thurston, "On the geometry and dynamics of diffeomorphisms of surfaces", *Bull. Amer. Math. Soc.* (*N.S.*) **19**:2 (1988), 417–431.

[Thurston 1997] W. P. Thurston, *Three-dimensional geometry and topology*, vol. 1, Princeton Mathematical Series **35**, Princeton Univ. Press, Princeton, NJ, 1997.

[Thurston 1998] W. P. Thurston, "Minimal stretch maps between hyperbolic surfaces", preprint, 1998. Available at www.arXiv.org/abs/math.GT/9801039.

[Tucker 1974] T. W. Tucker, "Non-compact 3-manifolds and the missing-boundary problem", *Topology* **13** (1974), 267–273.

[Tucker 1975] T. W. Tucker, "A correction to a paper of A. Marden (Ann. of Math. (2) **99** (1974), 383–462)", *Ann. of Math.* (2) **102**:3 (1975), 565–566.

[Tukia 1985a] P. Tukia, "Differentiability and rigidity of Möbius groups", *Invent. Math.* **82**:3 (1985), 557–578.

[Tukia 1985b] P. Tukia, "On isomorphisms of geometrically finite Möbius groups", *Inst. Hautes Études Sci. Publ. Math.* no. 61 (1985), 171–214.

[Tukia 1985c] P. Tukia, "Quasiconformal extension of quasiisometric mappings compatible with a Möbius group", *Acta Math.* **154** (1985), 153–193.

[Tukia 2005] P. Tukia, 2005. private communication.

[Van Vleck 1919] E. B. Van Vleck, "On the combination of non-loxodromic substitutions", *Trans. Amer. Math. Soc.* **20**:4 (1919), 299–312.

[Vuorinen 1988] M. Vuorinen, *Conformal geometry and quasiregular mappings*, Lecture Notes in Math. **1319**, Springer, Berlin, 1988.

[Wada 2003] M. Wada, "Exploring the space of twice punctured torus groups", Aug. 3 2003. Lecture at the Cambridge workshop *Spaces of Kleinian Groups and Hyperbolic* 3-*Manifolds*.

[Wada 2006] M. Wada, "OPTi's algorithm for discreteness determination", *Experiment. Math.* **15**:1 (2006), 61–66.

[Wada ≥ 2007a] M. Wada, "Cyclic Groups", Available at wada@ics.nara-wu.ac.jp. computer program to visualize the isometric polyhedra for cyclic loxodromic groups.

[Wada ≥ 2007b] M. Wada, "OPTi", Available at http://vivaldi.ics.nara-wu.ac.jp. computer program to visualize the isometric polyhedra for quasifuchsian once-punctured torus groups.

[Waldhausen 1968] F. Waldhausen, "On irreducible 3-manifolds which are sufficiently large", *Ann. of Math.* (2) **87** (1968), 56–88.

[Wang 1969] H.-c. Wang, "Discrete nilpotent subgroups of Lie groups", *J. Differential Geometry* **3** (1969), 481–492.

[Wang 1972] H. C. Wang, "Topics on totally discontinuous groups", pp. 459–487 in *Symmetric spaces* (St. Louis, MO, 1969–1970), edited by W. M. Boothby and G. L. Weiss, Pure and Appl. Math. **8**, Dekker, New York, 1972.

[Weeks 1993] J. R. Weeks, "Convex hulls and isometries of cusped hyperbolic 3-manifolds", *Topology Appl.* **52**:2 (1993), 127–149.

[Weeks 2004] J. Weeks, "The Poincaré dodecahedral space and the mystery of the missing fluctuations", *Notices Amer. Math. Soc.* **51**:6 (2004), 610–619.

[Weeks 2005] J. Weeks, "Computation of hyperbolic structures in knot theory", in *Handbook of knot theory*, edited by W. W. Menasco and M. B. Thistlethwaite, Elsevier, Amsterdam, 2005.

[Weeks n.d.] J. R. Weeks, "SnapPea", n.d. Available at http://www.geometrygames.org/SnapPea/index.html. software to compute hyperbolic manifolds.

[Weiss 2002] H. Weiss, *Local rigidity of* 3-*dimensional cone-manifolds*, Ph.D. thesis, 2002.

[Whitehead 1978] G. W. Whitehead, *Elements of homotopy theory*, Graduate Texts in Mathematics **61**, Springer, New York, 1978.

[Wielenberg 1977] N. J. Wielenberg, "Discrete Moebius groups: fundamental polyhedra and convergence", *Amer. J. Math.* **99**:4 (1977), 861–877.

[Wielenberg 1978] N. Wielenberg, "The structure of certain subgroups of the Picard group", *Math. Proc. Cambridge Philos. Soc.* **84**:3 (1978), 427–436.

[Wielenberg 1981] N. J. Wielenberg, "Hyperbolic 3-manifolds which share a fundamental polyhedron", pp. 505–513 in *Riemann surfaces and related topics* (Stony Brook, NY, 1978), edited by I. Kra and B. Maskit, Ann. of Math. Stud. **97**, Princeton Univ. Press, Princeton, NJ, 1981.

[Wright 2006] D. J. Wright, "Searching for the cusp", in *Spaces of kleinian groups*, edited by Y. Minsky et al., London Math. Soc. Lecture Notes **329**, Camb. Univ. Press, 2006.

Index