From Algebra to Geometry: A Hyperbolic Odyssey

The construction of geometric cone-manifold structures with prescribed holonomy

Daniel Virgil Mathews

Submitted in total fulfilment of the requirements of the degree of Master of Science
(by research)

September 2005

Department of Mathematics and Statistics
University of Melbourne

Produced on acid free paper.

Abstract

This thesis examines the relationship between geometric cone-manifold structures on surfaces, and algebraic representations of the fundamental group into a group of isometries. A geometric cone-manifold structure on a surface, with all interior cone angles being integer multiples of 2π, determines a holonomy representation of the fundamental group. We ask, conversely, when a representation of the fundamental group is the holonomy of a geometric cone-manifold structure. We consider 2-dimensional hyperbolic geometry and expand upon the known results.

We prove results for the punctured torus and higher genus surfaces. Our techniques are in the main low-powered, constructing fundamental domains for hyperbolic cone-manifold structures. We use various ideas to allow us to deduce facts about the geometry of a representation, from algebraically derivable data.

Central to these techniques are the Euler class of a representation, a geometric concept with an algebraic description, and the universal covering group $\widetilde{PSL_2\mathbb{R}}$ of the group of orientation-preserving isometries of $\mathbb{H}^2$. We also introduce the apparently new (though intuitive) geometric notion of the "twist" of a hyperbolic isometry, which is related to a purely algebraic notion of the "angle" of a matrix introduced by Milnor. We also require various simple results and constructions from hyperbolic geometry, and of course a basic understanding of the geometry of hyperbolic cone-manifolds.

For some of our constructions we need to change the basis of the fundamental group, and so we must consider the action of the outer automorphism group on the character variety. This action is measure-preserving with respect to a natural measure derived from a symplectic structure on the character variety. The action has interesting dynamics, being ergodic in certain regions. Putting these results together, and using various constructions in hyperbolic geometry, we can construct hyperbolic cone-manifold structures on surfaces with prescribed holonomy, in a wide range of cases.

Declaration

This is to certify that the thesis comprises only my original work except where indicated in the preface; due acknowledgment has been made in the text to all other material used; and the thesis is approximately 30,000 words in length, inclusive of footnotes, but exclusive of tables, maps, appendices and bibliography.

Acknowledgments

I would like to acknowledge the support of my supervisor Craig Hodgson, whose advice has been invaluable. I would also like to thank my family for their continual support and understanding throughout the preparation of this thesis. Thanks also to Joe Healy for help with formatting and typesetting.

Contents

List of Figures

Preface

The original parts of this thesis are primarily in sections 3.4, 3.6–3.7, chapter 5, sections 6.4–6.5 and chapters 7 and 8. The rest is introduction, exposition and discussion of well-known ideas and existing work (as cited in the bibliography and throughout the text) necessary for these results.

Chapter 1 is an introduction. Chapter 2 consists of geometric preliminaries, summarising some well-known ideas. Chapter 3 collates various results in hyperbolic geometry. Some of this is introduction, some is technical detail, with proofs derived from various sources as cited, and some is drawn from work of others, particularly Milnor. The notion of "twist" in section 3.4 is intuitive but appears to be new, as does the geometric interpretation of Milnor's function Θ. Chapter 4 is an introduction to various ideas regarding the Euler class and representation and character varieties, together with a description of work of Goldman regarding symplectic geometry. Chapter 5 discusses the geometry of punctured tori with hyperbolic cone-manifold structures, and although rather simple appears to be new. Chapter 6 is algebraic: the first three sections are a summary of known results; the fourth and fifth sections are elaborations of various relevant details. Chapter 7 is original, although the algorithm in 7.6.2 is the "opposite" of Goldman's algorithm in [30] (which is really just a greedy algorithm in any case). Chapter 8 is also mostly original. However, section 8.2 is a reproof of a known theorem using apparently new techniques, and relying essentially on a difficult theorem from [23]. And section 8.4 builds upon Goldman's work in [30].

Chapter 1

Introduction

1.1 The problem

The study of geometric structures on manifolds is an interesting and important endeavour, at least since the rise to prominence of the geometrisation programme of Thurston. The main questions of geometrisation are posed for three-manifolds. But some problems are difficult enough in two dimensions, or even more difficult, such as the one we deal with in this thesis.

It is well known that any geometric structure, i.e. an $(X, \operatorname{Isom} X)$ structure, on an orientable manifold M induces a holonomy representation of the fundamental group into the group of orientation-preserving isometries of the geometry X, that is, a homomorphism $\rho : \pi_1(M) \longrightarrow \operatorname{Isom} X$. This gives rise to connections between the algebra of the holonomy representation, and the geometry of the corresponding structure on the manifold.

From a hyperbolic structure, it is an easy matter to obtain the representation ρ. But the route from a representation ρ to a geometric structure is much longer and uphill. We ask: given a representation $\rho : \pi_1(M) \longrightarrow \operatorname{Isom} X$, is ρ the holonomy of a geometric structure on M? The question may be asked of any geometry, and is in general a difficult question.

The answer varies between different types of geometry, and depends on how broadly we define "geometric structure". We may ask for a complete geometric structure on a closed manifold. If our manifold has boundary, we may stipulate that the boundary be totally geodesic, or not. We may allow our boundary to have singularities such as "corners". And we may allow singularities of certain types inside our manifold. If we allow cone singularities, we obtain a *cone-manifold structure*. A representation ρ can only make sense as a holonomy of a cone-manifold structure, however, if every interior cone point has a cone angle which is an integer multiple of 2π. This broadening of the notion of geometric structure is quite a natural one to make, and has been taken elsewhere ([39], [23]).

In this thesis we focus on 2-dimensional hyperbolic geometry and broaden the known results. But first it is worth recounting the known results for other types of geometry.

Three-dimensional hyperbolic, Euclidean, spherical geometry; manifold with boundary. Let M be a 3-manifold with nonempty boundary, and let $(X, \operatorname{Isom} X)$ denote 3-dimensional hyperbolic, spherical or Euclidean geometry. In [39] Leleu proved that a representation $\rho : \pi_1(M) \longrightarrow \operatorname{Isom} X$ is the holonomy of an $(X, \operatorname{Isom} X)$-structure on M if and only if ρ lifts to the universal covering group $\widetilde{\operatorname{Isom} X}$. No cone points are required. However the boundary need not be totally geodesic, and will in general be complicated.

Three-dimensional hyperbolic geometry; cusped manifold; cone-manifold structures. Recall that a finite volume complete orientable hyperbolic 3-manifold M is diffeomorphic to the interior of a compact 3-manifold $\bar{M}$ whose boundary consists of tori. There are various ways the question has been attacked in this case. In [47] I investigated the representation varieties of simple hyperbolic knot complements $S^3 - K$ via the A-polynomial $A_K(x, y)$ of K. The A-polynomial encodes information about the restriction of a representation $\pi_1(S^3 - K) \longrightarrow \operatorname{Isom}^+ \mathbb{H}^3$ to the peripheral subgroup (see [11]). I found in several examples that each branch of the variety defined by $A_K(x, y)$ had a geometric interpretation describing the holonomy of hyperbolic cone-manifold structures on $S^3 - K$. It is known ([36], [47]) that this is true for twist knots.

In other directions, one can show that there is a well-defined "volume" associated to a representation $\pi_1(M) \longrightarrow PSL_2\mathbb{C}$, and that it is maximised at the representation of the unique complete hyperbolic structure — it is unique by Mostow rigidity. See, e.g. [59], [16], [20].

Two-dimensional complex projective geometry; closed surface; cone-manifold structures. For an oriented closed surface S, Gallo, Kapovich and Marden proved in [23] that a representation $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{C}$ is the holonomy of a complex projective cone-manifold structure if and only if ρ is nonelementary. If ρ lifts to a representation into $SL_2\mathbb{C}$, then a complete complex projective structure is possible. Otherwise a single cone point of angle 4π is sufficient.

From this point forward, however, we narrow our focus to two-dimensional hyperbolic geometry. The problem is very interesting, and seemingly more complicated, in this case.

Complete hyperbolic structures with totally geodesic boundary. For such structures, the question was answered by Goldman in his thesis [25]. For a closed surface S with $\chi(S) < 0$, a representation $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ determines an *Euler class* $\mathcal{E}(\rho)$. The Euler class is a 2-dimensional cohomology class on S, hence a multiple of the fundamental class. The Euler class may be any multiple of the fundamental class between $\chi(S)$ and $-\chi(S)$, and it parametrises the connected components of the representation space ([28]). Goldman proved that ρ is the holonomy of a hyperbolic structure on S if and only if the Euler class is $\pm\chi(S)$ times the fundamental cohomology class. If S has boundary, then the same machinery applies, and the same theorem holds, provided that each boundary curve is sent to a non-elliptic isometry. In this case we obtain a *relative* Euler class. We will discuss these ideas in more detail, and in fact prove Goldman's theorem, in the sequel.

It is worth mentioning that all the answers known to the fundamental question "when is ρ the holonomy of a geometric structure?" involve lifting ρ to the universal cover of the isometry group. The Euler class can be thought of as an *obstruction* to lifting ρ to a map $\pi_1(S) \longrightarrow \widetilde{PSL_2\mathbb{R}}$. So from Goldman's result, a holonomy representation does not lift to $\widetilde{PSL_2\mathbb{R}}$; in fact, such representations are as "unliftable" as possible!

Cone-manifold structures. The question then arises: for a surface S with $\chi(S) < 0$, do the other components of the space of representations have any reasonable geometric interpretation? It is natural to consider hyperbolic cone-manifold structures on our manifold, provided that all interior cone points have a cone angle which is an integer multiple of 2π. It is easy to prove, as we will see, that if ρ is the holonomy of such a cone-manifold structure on S, then the Euler class is related in a simple way to the number and type of cone points. If S has boundary, then we may require that the boundary be totally geodesic, or piecewise geodesic with a small number of corners. Allowing arbitrarily many corners rapidly trivialises the problem, permitting us to construct geometric structures easily.

Other possibilities. We might also allow folding of our hyperbolic structure, but then we must restrict the number of folds very tightly to avoid trivialising the problem: see e.g. [58]. Another way to broaden the question is to relinquish control over the boundary of a surface, and not requiring it to be totally or even piecewise geodesic. Then the answer is simple, but the boundary may be very complicated. It seems that allowing cone-type singularities is the most reasonable broadening of the notion of a hyperbolic structure.

1.2 Synopsis

In this thesis we prove several theorems regarding hyperbolic cone-manifold structures on surfaces. For a punctured torus, we prove that apart from a small family of virtually abelian representations, *any* representation is the holonomy of a cone-manifold structure, where we allow at most one corner on the boundary. We give a seemingly new low-powered proof of Goldman's theorem that the representations of a surface corresponding to complete hyperbolic structures are those with extremal Euler class. We prove a theorem regarding the genus 2 closed surface: we show that under certain circumstances we can construct a cone-manifold structure from a representation with Euler class ± 1. And we show that for a closed surface *almost* any representation with Euler class $\pm(\chi(S)+1)$ (i.e. "one off" from extremal), which sends some non-separating simple closed curve to an elliptic element of $PSL_2\mathbb{R}$, is the holonomy of a cone-manifold structure with at most one cone point with angle 4π. The word "almost" involves introducing a measure on a suitable space.

There are many questions still awaiting answers. Note the partial nature of these results; the questions are difficult. In [58], Tan gives an example of a representation with Euler class 2 on a closed genus 3 surface which is not the holonomy of any hyperbolic cone-manifold structure. He reported that Goldman and Neumann had a proof that for a surface S with $\chi(S) < 0$, every representation with Euler class $\pm(\chi(S)+1)$ is the holonomy of a hyperbolic cone-manifold structure on S with one interior cone point of angle 4π. But the status of this proof, which remains unpublished, is unclear. We can still ask: for representations of Euler class $\pm 1, \pm 2, \ldots, \pm(\chi(S)+1)$, is the set of representations which are holonomies of hyperbolic cone-manifold structures dense in the set of representations of that Euler class? Using ergodicity methods one might hope that the set of such representations is conull, i.e. almost every representation is a holonomy representation in these cases.

This thesis is organised as follows. On the one hand, we need to understand the basic geometry of cone-manifolds. On the other hand, we need to investigate the algebra of representations of a surface group into $PSL_2\mathbb{R}$. And in between, we need to develop some of the wonderful properties, algebraic and geometric, of the universal covering group $\widetilde{PSL_2\mathbb{R}}$. To construct cone-manifolds from mere representations, we will need a swag of tricks up our sleeve. Given a representation, we need to tell as much about the geometry as we can, simply by looking at a few algebraic parameters. Given a geometric arrangement of isometries, we need shortcuts to construct a cone-manifold structure. Once we have established all this, we will commence our investigations of the punctured torus, the genus 2 surface and higher genus surfaces.

In chapter 2 we present a brief introduction to geometric structures on manifolds, and cone-manifolds. In particular we consider the properties of geodesics in cone-manifolds, and the nature of cone singularities.

In chapter 3 we develop some results in hyperbolic geometry that are required in the sequel. Most of this is just a summary of known results, but some of it is a reworking of old ideas and some of it appears to be new. We establish a number of results connecting the arrangement of isometries in the hyperbolic plane to algebraic properties of matrices in $SL_2\mathbb{R}$. We examine the group $\widetilde{PSL_2}\mathbb{R}$. We introduce the notion of the "twist" of a hyperbolic isometry at a point, which is a highly intuitive geometric concept, but I have not seen it developed explicitly before. We will consider a similar algebraic notion of "angle" associated to a matrix, introduced by Milnor in [49]. We will give a geometric interpretation to Milnor's function in terms of hyperbolic geometry and "twisting". And we will examine how information can be extracted from the traces of matrices.

In chapter 4 we introduce the Euler class, the cohomology class associated to a representation ρ. We give a brief description and establish the primary method for its calculation, via the algebra of $\widetilde{PSL_2}\mathbb{R}$. We also examine the space of representations and the space of characters, which are objects from algebraic geometry; and we introduce a symplectic structure, and a measure on the character variety, which we shall use in our "almost every" statement.

In chapter 5 we analyse the geometry of punctured tori with hyperbolic cone-manifold structures with one corner point. We will show how they can be decomposed into a fundamental domain. Conversely, we will give a simple method for constructing such a fundamental domain, and hence a hyperbolic cone-manifold structure with prescribed holonomy. We simply have to find a certain pentagon to bound an immersed open disc in $\mathbb{H}^2$. We will examine the lack of rigidity in these geometric structures — one representation can be the holonomy for a large family of geometric structures, with many pictures. This chapter defers proofs and offers an intuitive explanation of what is happening in the underlying geometry. Finally we will establish a relationship between our notion of "twist" and the corner angle which arises. Putting together our algebraic and geometric results gives us significant tools with which to analyse representations. They are not particularly high-powered techniques, but they allow significant insight. Indeed, our results relating twisting to corner angles are really proved just by following vectors around a picture; these methods can really be thought of as quite naive and rather kindergarten-like.

In chapter 6 we discuss some of the remarkable algebraic properties of representations from the free group on two generators, which is the fundamental group of the punctured torus, to $PSL_2\mathbb{R}$. Nielsen's theorem shows just how closely algebra and geometry are related in this case. The character variety has a very simple description, and changes of basis in the fundamental group have a simple description in terms of Markoff moves. We will characterise the virtually abelian representations in terms of the character variety.

In chapter 7 we prove our main result on punctured tori, constructing hyperbolic cone-manifold structures with no interior cone points and at most one corner point, for all representations ρ except the virtually abelian ones. We classify the possible representations according to a single parameter which is natural in light of Nielsen's theorem. In the easier cases the precise geometric arrangement can be deduced immediately from our previous results and we can easily construct a fundamental domain. In the most difficult case, we must apply an algorithm to change basis in our fundamental group, using the Markoff moves described in chapter 6, until we can obtain a geometric arrangement from which construction of the requisite pentagonal fundamental domains is possible.

In chapter 8 we prove our results for higher genus surfaces. We apply our methods to give a proof of Goldman's theorem, which is perhaps more low-powered than the original proof, although we must rely on a theorem of Gallo, Marden and Kapovich in [23] guaranteeing decompositions along hyperbolic curves, for which the proof is long and detailed. We prove partial results allowing us to construct hyperbolic cone-manifold structures on the genus 2 surface, using our knowledge of the Euler class and $\widetilde{PSL_2}\mathbb{R}$ to classify the possible splittings of the surface into two punctured tori. Our task is to find not one but two pentagonal fundamental domains which fit together. And we prove our "almost every" theorem giving us hyperbolic cone-manifold structures for representations which have "one-off-extremal" Euler class and which send certain curves to elliptics. We use results of Goldman [30], that the action on the character variety induced by changes of basis is sometimes *ergodic* with respect to a certain measure. These allow us to change basis to "almost go almost anywhere" in the character variety, and hence, simply by changing basis, alter the geometric situation almost entirely as we please! This is a very interesting technique which we hope will have applications to other results.

At the start of their paper solving the equivalent problem for complex projective geometry, Gallo, Kapovich and Marden described their efforts on the problem as follows: "We three authors decided to join together to pool the fruits of a decade of our individual and collaborative research relating to the main result." That may indicate the difficulty of the question, especially as the complete answer in the 2-

dimensional hyperbolic case is certainly more complicated. But undaunted we will plunge into the breach and attempt to shed some light on this riddle.

1.3 Main results

We prove the following results.

Theorem A *Let S be a punctured torus. A representation $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ is the holonomy of a hyperbolic cone-manifold structure on S with at most one corner point and no interior cone points if and only if ρ is not virtually abelian.*

We also reprove the theorem of Goldman [25]:

Goldman's Theorem *Let S be an orientable surface with $\chi(S) < 0$, and let ρ be a representation $\pi_1(S) \longrightarrow PSL_2\mathbb{R}$. If S has boundary, assume ρ takes each boundary curve to a non-elliptic element, so the relative Euler class $\mathcal{E}(\rho)$ is well-defined. A representation ρ is the holonomy of a complete hyperbolic structure on S with totally geodesic or cusped boundary components (respectively as the boundary curve is hyperbolic or parabolic) if and only if $\mathcal{E}(\rho)[S] = \pm\chi(S)$.*

In the case of a closed surface, this becomes:

Corollary *Let S be a closed orientable surface of genus $g \geq 2$. A representation $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ is the holonomy of a complete hyperbolic structure on S if and only if $\mathcal{E}(\rho)[S] = \pm\chi(S)$*

In the case of a genus 2 surface, we have the following result.

Theorem B *Let S be a genus 2 closed surface. Let $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ be a representation with $\mathcal{E}(\rho)[S] = \pm 1$. Suppose that there is a separating curve C on S such that $\rho(C)$ is not hyperbolic. Then ρ is the holonomy of a hyperbolic cone-manifold structure on S with one cone point of angle 4π.*

And for a general closed surface, we have the following.

Theorem C *Let S be a closed orientable surface of genus $g \geq 2$. Almost every representation $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ with $\mathcal{E}(\rho)[S] = \pm(\chi(S) + 1)$, which sends some non-separating simple closed curve C to an elliptic, is the holonomy of a hyperbolic cone-manifold structure on S with a single cone point with cone angle 4π.*

As we proceed, we will need to introduce a measure on some space of representations (namely, the *character variety*) so that this statement makes sense.

Chapter 2

Preliminaries

Throughout, let S be an orientable surface.

2.1 Geometric structures on manifolds

Recall that a geometric structure on S is a metric on S so that every point of S has a neighbourhood isometric to a standard ball neighbourhood in our model geometry X.

Equivalently, because of this local equivalence with the model geometry, we can consider a geometric structure via a maximal atlas. Letting U be an open set in S, recall that a *coordinate chart* is a map $\phi : U \longrightarrow X$ which is a homeomorphism onto its image. If we have two coordinate charts ϕ_i, ϕ_j defined on open sets U_i, U_j which overlap, we would like the two coordinate charts to be compatible. Thus the *transition map* $\gamma_{ij} = \phi_i \circ \phi_j^{-1}$ is required to be (a restriction of) an isometry of X. A set of compatible coordinate charts whose domains cover S is called an *atlas*. An atlas then gives an explicit description of the local correspondence between S and X, but clearly different atlases can correspond to equivalent geometric situations (for instance, if one atlas is a subset of another). In order to have a unique definition of a geometric structure, we can define a geometric structure as a *maximal atlas* on S.

A geometric structure gives rise to a *developing map*. Commencing with a particular coordinate chart $\phi_i : U_i \longrightarrow X$, we have a map of a local (contractible) neighbourhood of a point into X. If we wander further afield in S, we may analytically continue this map. For instance, if $U_i \cap U_j \neq \emptyset$ then replacing ϕ_j with $\phi'_j = \gamma_{ij} \circ \phi_j$ gives us two charts which agree on $U_i \cap U_j$. Continuing such a process for all paths in S, i.e. on the *universal cover* $\tilde{S}$, we obtain the developing map $\mathcal{D} : \tilde{S} \longrightarrow X$. From a particular geometric structure, the developing map is well-defined up to the choice of where to start and which direction to go — i.e., up to conjugation by isometries of X.

Walking around a loop C in S, if the loop is trivial then under our developing map $\mathcal{D}$ we traverse a loop in X starting at a point x_0. But if C is homotopically non-trivial, then we end up passing through several local neighbourhoods $U_0, U_1, \ldots, U_n = U_0$ with coordinate charts $\phi_0, \phi_1, \ldots, \phi_n = \phi_0$ and undergoing several coordinate changes $\gamma_{0,1}, \gamma_{1,2}, \ldots, \gamma_{n-1,n}$ which are isometries of X. Thus the image of C under $\mathcal{D}$ ends at the point $\gamma_{n-1,n} \circ \gamma_{n-2,n-1} \circ \cdots \circ \gamma_{0,1} x_0$. This isometry, which depends only on the homotopy class of C and describes the action within X as we walk along the developing image of this curve, is called the *holonomy* of the curve C. Thus we obtain the *holonomy map* or *holonomy representation* $\rho : \pi_1(S) \longrightarrow \operatorname{Isom} X$. Note that $\pi_1(S)$ acts on $\tilde{S}$ by deck transformations and, via ρ, acts on X via isometries. This action is equivariant with respect to the developing map $\mathcal{D}$, i.e. if T_α is the deck transformation corresponding to α, then $\mathcal{D} \circ T_\alpha = \rho(\alpha) \circ \mathcal{D}$. Now for $\alpha, \beta \in \pi_1(S)$ we have $T_\alpha \circ T_\beta = T_{\alpha\beta}$, provided that composition in $\pi_1(S)$ is written *left to right* and composition of functions is (as usual!) written *right to left*: this is an important convention to remember! Then

$$\rho(\alpha\beta) \circ \mathcal{D} = \mathcal{D} \circ T_{\alpha\beta} = \mathcal{D} \circ T_\alpha \circ T_\beta = \rho(\alpha) \circ \mathcal{D} \circ T_\beta = \rho(\alpha)\rho(\beta) \circ \mathcal{D}$$

so ρ is a homomorphism of groups.

A geometric structure can also be described as a type of section of a certain fibre bundle. This description also offers an approach to the question in which we are interested: given a homomorphism $\rho : \pi_1(S) \longrightarrow \operatorname{Isom} X$, is ρ the holonomy of a geometric structure on S? Finding such a structure corresponds to finding an appropriate section of this fibre bundle, which we describe now.

We define the bundle $\mathcal{F}(S, X, \rho)$ as the quotient of $\tilde{S} \times X$ by $\pi_1(S)$. Here $\pi_1(S)$ acts on $\tilde{S}$ by deck transformations, and acts on X via the isometries given by ρ. The quotient by this action is a flat X-bundle over S with holonomy ρ. The product $\tilde{S} \times X$ is foliated by lines of the form $\tilde{S} \times \{x\}$, for each individual $x \in X$. This foliation then descends to a foliation on $\mathcal{F}(S, X, \rho)$. A section of the bundle transverse to this foliation is precisely a geometric structure on S. The section $s : S \longrightarrow \mathcal{F}(S, X, \rho)$ immediately gives a developing map by lifting to a map $\tilde{s} : \tilde{S} \longrightarrow \tilde{S} \times X$, which is equivariant under the action of $\pi_1(S)$, and projecting onto the second coordinate.

The transversality condition described above can be reformulated in terms of the developing map. Transversality of s to the foliation is equivalent to $\tilde{s}$ always having nonzero derivative in the X coordinate. If we project $\tilde{s}$ onto the second coordinate to obtain the developing map, the requirement is then that the developing map must have nowhere zero Jacobian. That is, the developing map must be an immersion.

2.2 Cone-manifolds

We now give an introduction to the notion of cone-manifolds. For our purposes we only need to define hyperbolic cone-manifolds, though the following definition has obvious generalisations to other geometries.

Definition 2.2.1 (Hyperbolic cone-manifold) *An n-manifold S with a metric is a* hyperbolic cone manifold *if there exists a triangulation of S such that*

(i) the link of each simplex is piecewise linear homeomorphic to the n-sphere, and

(ii) the restriction of the metric to each simplex is isometric to a geodesic simplex in hyperbolic space.

Recall that the *open star* of a point $x \in S$ in a triangulated manifold S is the union of the interiors of the simplices of S containing x. The *open star* of a simplex σ is the union of the interiors of the simplices of S containing σ. The *closed star* is defined similarly but without taking the interiors of the simplices. For each simplex τ_i containing σ, let σ_i denote the simplex opposite σ in τ_i. The *link of the simplex σ*, denoted $\mathrm{Lk}(\sigma, S)$, is the union of the σ_i. About a point x there is also the *geometric link of the point x*, denoted $\mathrm{Lk}(x, S)$, which is the set of unit vectors at x pointing into S.

We only need consider cone-manifolds in dimension 2. A hyperbolic cone-manifold S is then simply a surface obtained by piecing together geodesic triangles in $\mathbb{H}^2$. Points p in the interior of S have neighbourhoods locally isometric to $\mathbb{H}^2$, except possibly at some vertices of the triangulation, around which the angles sum to $\theta \neq 2\pi$. Such points are called (interior) *cone points*. The neighbourhood of such a cone point is isometric to a wedge of angle θ in $\mathbb{H}^2$, with sides glued (i.e. a cone). The angle θ is called the *cone angle* at p. Letting $\theta = 2\pi(1 + s)$, we call the number s the *order* of the cone point, following [61]. If S has boundary then this boundary will be piecewise geodesic. There may be vertices on the boundary around which the angles sum to $\theta \neq \pi$. Such a point is called a *corner point* and θ is the *corner angle*. Letting $\theta = 2\pi(\frac{1}{2} + s)$, then s is the *order* of the corner point. A corner point has a neighbourhood isometric to a wedge of angle θ in $\mathbb{H}^2$ (without sides glued). A *singular point* is a cone or corner point. The set of singular points is called the *singular locus*. In general the singular locus of an n-dimensional cone-manifold is a union of totally geodesic closed simplices of dimension $n - 2$. Other points are called *regular points*.

Note a cone or corner angle can be any positive real number — it can be more than 2π. We will be dealing with many large cone angles, and it is important to have some understanding of their properties.

Consider the following examples.

(i) An n-orbifold is a cone-manifold. At each point of the singular locus the cone angle is of the form $2\pi/k$ for some $k \in \mathbb{N}$.

(ii) A branched cover of a hyperbolic n-manifold over a piecewise geodesic branching set is a cone-manifold with all cone angles of the form $2\pi k$ for some $k \in \mathbb{N}$.

(iii) Many hyperbolic cone 3-manifolds arise in performing hyperbolic Dehn surgery on cusped hyperbolic 3-manifolds. For instance, $(p, 0)$ surgery on a manifold with a single cusp gives rise to a hyperbolic cone-manifold with singular locus a circle with cone angle $2\pi/p$.

(iv) A hyperbolic cone-manifold structure on a genus g surface can be obtained by taking a regular $4g$-gon in $\mathbb{H}^2$ and identifying pairs of edges in a standard way. The interior angles in the $4g$-gon can take any value from 0 to the Euclidean angle of $\pi(2g-1)/2g$. For each angle θ we obtain a hyperbolic cone-manifold structure with a single cone point of angle $4g\theta$, which may vary anywhere strictly between 0 and $2\pi(2g-1)$.

In this last example we see that there are limits on the allowable cone and corner angles in a 2-dimensional hyperbolic cone-manifold. These limits can be deduced directly from the Gauss-Bonnet theorem. Recall that, for a totally geodesic hyperbolic triangle Δ with interior angles α, β, γ, we have

$$\int_\Delta K \, dA = \alpha + \beta + \gamma - \pi$$

where K is the Gaussian curvature, $K = -1$ in the hyperbolic plane. We take a triangulation of S as in the definition above, so that all cone points and corner points occur as vertices. We now sum over all the triangles in the triangulation. Let there be V vertices, E edges and F faces, so $\chi(S) = V - E + F$. We obtain

$$\begin{aligned} 0 > \int_S K \, dA &= \sum_{\text{Interior } v_i} 2\pi(1 + s_i) + \sum_{\text{Boundary } v_i} 2\pi\left(\frac{1}{2} + s_i\right) - \pi F \\ &= 2\pi V_{\text{interior}} + \pi V_{\text{boundary}} - \pi F + 2\pi \sum s_i \\ &= 2\pi\chi(S) + 2\pi \sum s_i. \end{aligned}$$

The first equality follows from grouping the angles around distinct vertices and using the definition of order of a cone point (note at vertices which are not cone or corner points, $s_i = 0$). The second equality is clear. The last equality follows from $2E = 3F + V_{\text{boundary}}$ and $V - E + F = \chi(S)$. We then obtain

Lemma 2.2.2 *Let S be a surface (with or without boundary). A hyperbolic cone-manifold structure on S with cone and corner points having orders s_i satisfies*

$$\sum s_i < -\chi(S).$$

■

This gives us a bound on the possible excess curvature that can be present in cone points.

Note that in any hyperbolic cone-manifold, a sufficiently small loop around an interior cone point v is homotopically trivial. Therefore, if a holonomy map is to be well-defined, the corresponding isometry of $\mathbb{H}^2$ must be the identity. This can only occur if the cone angle is an integer multiple of 2π, in which case a loop about v, under our developing map, winds around some $\mathcal{D}(\tilde{v})$ a number of times but forms a closed loop. However there is no such problem with corner points, which *a priori* may have any corner angle, subject to the bounds discussed above.

We state without proof some of the basic properties of curves, lengths and geodesics on hyperbolic cone-manifolds. The basic reference here is [8].

First recall some notions which may be defined in any metric space. In any metric space (S, d), a curve (i.e. continuous map) $C : [0,1] \longrightarrow S$ has *length* given by

$$\sup_{a=t_0\leq t_1\leq\cdots\leq t_n=b} \sum_{i=0}^{n-1} d(c(t_i), c(t_{i+1})),$$

where the supremum is taken over all possible partitions $a = t_0 \leq t_1 \leq \cdots \leq t_n = b$. A curve with finite length is called *rectifiable.* A *unit speed geodesic* is a curve $C : [a,b] \longrightarrow S$ such that for all $t \in [a,b]$ and t' sufficiently close to t, we have $d(C(t), C(t')) = |t - t'|$. A curve which is a reparametrisation of a unit speed geodesic is called a *geodesic.* A metric space in which every two points are joined by a geodesic is called a *geodesic space.* If the distance $d(x,y)$ between every pair of points $x, y \in M$ is equal to the infimum of the length of rectifiable curves between them, then d is called a *length metric* and (S, d) a *length space.* Not every metric is a length metric, though we may always define an induced length metric [8, I.3].

Let S be a hyperbolic cone-manifold. We may define a metric on S in one of two ways. We may consider a Riemannian metric on S, which is locally isometric to $\mathbb{H}^2$ at regular points. If we can also describe a neighbourhood of the singular locus via a (singular) Riemannian metric, then we may define a metric by letting $d(x,y)$ be the infimum of the Riemannian length of all piecewise differentiable paths beteen x and y. Alternatively, since S is triangulated, each simplex has a metric inherited from hyperbolic space. We can combine these to find a global metric. There are several ways to do this; methods using *strings* and *chains* are described in [8]. Not surprisingly, these two approaches yield the same resulting metric, which is a length metric.

We may define a standard hyperbolic cone neighbourhood, on which a cone point may be modelled. Given any *spherical* $(n-1)$-manifold M, the n-dimensional *hyperbolic open cone* $\text{Cone}(M,R)$ of radius R based on M is $M \times [0,R)$ with $M \times \{0\}$ collapsed to a point. We write $(m,t) \in \text{Cone}(M,R)$ as mt; we think of points in M being multiplied by a scalar factor. The equivalence class of the point $(m,0)$ is called the *vertex* of the cone. In the case $n=2$ the $(n-1)$-dimensional spherical cone-manifold M will simply be either a round circle or arc. If we let $[0,R)$ and M have Riemannian metrics dr and $d\theta$ respectively, then we can define infinitesimal distance on $\text{Cone}(M,R)$ as

$$ds^2 = dr^2 + \sinh^2 r d\theta^2,$$

this being the standard form for distance in $\mathbb{H}^n$ in polar coordinates.

A metric on $\text{Cone}(M,R)$ can be defined by integrating this expression. This metric has an explicit description. Let d_M^π be the metric on the spherical cone-manifold M defined by "truncation":

$$d_M^\pi(x,y) = \min(\pi, d_M(x,y)).$$

Then for $x_1, x_2 \in \text{Cone}(M,R)$, $x_1 = t_1 m_1$, $x_2 = t_2 m_2$, we define $d(x_1,x_2)$ to be the nonnegative number satisfying

$$\cosh\left(d(x_1,x_2)\right) = \cosh t_1 \cosh t_2 - \sinh t_1 \sinh t_2 \cos\left(d_M^\pi(m_1,m_2)\right).$$

One can prove that this function d is a metric on $\text{Cone}(M,R)$. One can also prove that the neighbourhood of any singular point $x \in S$ is of the form $\text{Cone}(M,R)$ for some M; in fact, $M = \text{Lk}(x,S)$. See [8, I.5] for details. For a regular point p in the interior of a 2-dimensional hyperbolic cone-manifold S, we have $\text{Lk}(x,S) = S^1$, the round circle, and $\text{Cone}(S^1,R)$ is isometric to a ball in $\mathbb{H}^2$.

The expression for the metric is somewhat bizarre. Note that for points x_1, x_2 separated by sufficiently small angles, the metric is just that of hyperbolic geometry, given by the hyperbolic cosine rule. In general, a geodesic C through the vertex x, after reparametrisation, has the form

$$C : (-\epsilon,\epsilon) \longrightarrow \text{Cone}(\text{Lk}(x,S),\epsilon), \quad C(t) = \begin{cases} -mt, & t<0, \\ m't, & t>0 \end{cases}.$$

Then for small $\delta > 0$ we have

$$\cosh d(C(-\delta), C(\delta)) = \cosh^2\delta - \sinh^2\delta \cos\left(d_{\text{Lk}(x,S)}^\pi(m,m')\right),$$

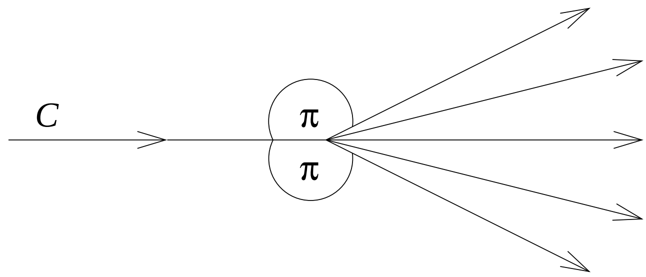

Figure 2.1: Extensions of a geodesic segment at a large angle cone point.

but for C to be a geodesic we must have $d(C(-\delta), C(\delta)) = 2\delta$ and hence

$$\cosh 2\delta = \cosh^2 \delta - \sinh^2 \delta \cos d^{\pi}_{\mathrm{Lk}(x,S)}(m, m').$$

It follows that, if C is a geodesic, then $d^{\pi}_{\mathrm{Lk}(x,S)}(m, m') = \pi$, i.e. $d_{\mathrm{Lk}(x,S)}(m, m') \geq \pi$. Conversely, if this inequality is satisfied then C is a geodesic. Thus C is a geodesic if and only if it makes an angle of at least π at x.

At a regular point the condition that a geodesic must make an angle of π is well known! It follows that a cone point with cone angle less than 2π cannot lie on the interior of a geodesic: all geodesics avoid cone points with small cone angles. However, there are many geodesics through a cone point x with cone angle more than 2π: any two geodesics with an endpoint at x, making an angle of at least π on both sides, join together to form a single geodesic. Thus, unlike the situation at regular points, a geodesic segment with an endpoint at x can be continued in infinitely many directions! See figure 2.1.

Note this argument applies equally if x is an interior cone point, or a corner point of S.

Finally, a hyperbolic cone-manifold is a geodesic space: this is the Hopf-Rinow theorem. Indeed the version we prove applies to all complete locally compact length spaces. We will give a proof, based on [8], since we will need to use the methods again in section 5.1. We need a version of the Arzelà-Ascoli theorem. Recall that if X, Y are metric spaces with metrics d_X, d_Y, then a function $\gamma : X \longrightarrow Y$ is called *L-Lipschitz* if for all $a, b \in X$,

$$d_Y(\gamma(a), \gamma(b)) \leq L\, d_X(a, b).$$

Theorem 2.2.3 (Arzelà-Ascoli) *Suppose X, Y are metric spaces such that X has a countable dense subset D and Y is compact. Let $\gamma_1, \gamma_2, \ldots$ be a sequence of L-Lipschitz functions $X \longrightarrow Y$. Then there is a subsequence of $\gamma_1, \gamma_2, \ldots$ which converges uniformly on compact sets to an L-Lipschitz function $\gamma : X \longrightarrow Y$.* ∎

Theorem 2.2.4 (Hopf-Rinow) *Let X be a complete, locally compact length space. Then every closed bounded subset of X is compact. For any $x, y \in X$ there exists a shortest geodesic C between them, i.e. with $l(C) = d(x, y)$. In particular, X is a geodesic space.*

Proof For the first assertion we refer to [8, prop. I.3.7]. We prove the second assertion: given $x, y \in X$ we find a shortest geodesic. Since the metric d on X is a length metric, we may find rectifiable curves $\gamma_n : [0, 1] \longrightarrow X$ from x to y, parameterised proportional to arc length, such that $l(\gamma_n) < d(x, y) + 1/n$.

We wish to invoke the Arzelà-Ascoli theorem. First we check that γ_n is L-Lipschitz for some L, independent of n. Such L must satisfy for all $t_1, t_2 \in [0, 1]$, $d(\gamma_n(t_1), \gamma_n(t_2)) \leq L\,|t_1 - t_2|$. Note that, as γ is parameterised at constant speed over $[0, 1]$, $|t_1 - t_2|$ is the fraction of the curve γ_n we are considering. So

$$|t_1 - t_2| = \frac{l\left(\gamma_n|_{[t_1,t_2]}\right)}{l(\gamma_n)} > \frac{d(\gamma_n(t_1), \gamma_n(t_2))}{d(x, y) + 1}.$$

This inequality follows since $l(\gamma_n|_{[t_1,t_2]}) \geq d(\gamma_n(t_1), \gamma_n(t_2))$ (by definition of length), and also since $l(\gamma_n) < d(x, y)+1/n \leq d(x, y)+1$. Thus we may take $L = d(x, y)+1$.

Second we check that γ_n can be taken to map into a compact set. For this take (say) the closed ball of radius $d(x, y) + 1$ about x. From the first assertion this is compact. We now apply the Arzelà-Ascoli theorem and conclude that a subsequence of the γ_n converge uniformly to a path $\gamma : [0, 1] \longrightarrow X$ from x to y.

We now show that γ is a geodesic. For this it is enough to show $l(\gamma) = d(x, y)$. By definition of a length metric, $l(\gamma) \geq d(x, y)$. But being the limit of the γ_n which satisfy $l(\gamma_n) < d(x, y) + 1/n$, we have $l(\gamma) \leq d(x, y)$. This concludes the proof. ∎

Chapter 3

Some hyperbolic geometry

3.1 The hyperbolic plane

In this chapter we discuss a number of aspects of 2-dimensional hyperbolic geometry we will need in our results. Recall that the hyperbolic plane $\mathbb{H}^2$ can be described via the *upper half plane model*, which is the set of all complex numbers with positive imaginary part, with Riemannian metric $ds^2 = (dx^2 + dy^2)/y^2$. The group of orientation-preserving isometries of $\mathbb{H}^2$ is naturally isomorphic to $PSL_2\mathbb{R}$, the group of 2×2 real matrices with determinant 1 subject to the equivalence relation $A \sim -A$. We can denote elements of $PSL_2\mathbb{R}$ by matrices up to sign, and then the matrix

$$\pm \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

acts on the upper half plane model via the Möbius transformation

$$z \mapsto \frac{az + b}{cz + d}.$$

This is an orientation-preserving isometry of $\mathbb{H}^2$ and all isometries are of this form. Accordingly, we identify matrices in $PSL_2\mathbb{R}$ with isometries, in a standard abuse of notation. The *circle at infinity* in this model is $\mathbb{R} \cup \{\infty\}$. The geodesics in this model are the vertical Euclidean lines and the Euclidean circles perpendicular to the real axis.

There is also the *unit disc model*, which is the set of all complex numbers of magnitude less than 1, with Riemannian metric $ds^2 = 4(dx^2 + dy^2)/(1 - |z|^2)^2$. In this model the circle at infinity is the unit circle in the complex plane, and the geodesics are arcs of Euclidean circles perpendicular to the circle at infinity.

Non-trivial orientation-preserving isometries of $\mathbb{H}^2$ come in three types. The *hyperbolic* isometries $g \in PSL_2\mathbb{R}$ translate along an axis, denoted $\operatorname{Axis} g$, by a fixed distance d_g. Such an isometry is represented by a matrix g with $|\operatorname{Tr} g| > 2$, where $\operatorname{Tr} g = 2\cosh(d_g/2)$, and fixes two points at infinity. A *parabolic* isometry g fixes a single point at infinity, translates along horocycles about this point, and has $|\operatorname{Tr} g| = \pm 2$. An *elliptic* isometry fixes a point $q \in \mathbb{H}^2$ and rotates about it by some angle θ. It is represented by a matrix with $|\operatorname{Tr} g| < 2$, and $\operatorname{Tr} g = 2\cos(\theta/2)$.

Fixing an arbitrary unit tangent vector at an arbitrary basepoint in $\mathbb{H}^2$, we see that a hyperbolic isometry is uniquely determined by where it takes this unit tangent vector. Thus we may identify the unit tangent bundle of the hyperbolic plane $UT\mathbb{H}^2$, which is homeomorphic to $\mathbb{R}^2 \times S^1$, with the isometry group $PSL_2\mathbb{R}$.

Recall that for any oriented line l in $\mathbb{H}^2$, we may define *Fermi coordinates* with respect to l. Choose a point q on l as our basepoint. Any point $p \in \mathbb{H}^2$ is now given coordinates (x, h) where x denotes "distance along l" and h denotes "height above l". More precisely, from p we drop a perpendicular to meet l at p'. Then x is the signed distance from q to p', and h is the signed length of the perpendicular dropped. In this way the hyperbolic plane is identified with $\mathbb{R}^2$. The distance between $p_1 = (x_1, h_1)$ and $p_2 = (x_2, h_2)$ is then given by the following relation (see e.g. [10] p. 38):

$$\cosh d(p_1, p_2) = \cosh h_1 \; \cosh h_2 \; \cosh(x_2 - x_1) - \sinh h_1 \; \sinh h_2. \qquad (3.1)$$

3.2 Compositions of Isometries

We now prove a few results about the effect of composing several isometries, which we shall need later. Here as throughout, isometries or matrices will be multiplied from right to left.

Lemma 3.2.1 *Let $g, h \in PSL_2\mathbb{R}$ be two hyperbolic isometries with the same translation distance, whose axes intersect at a point r. Let $p = g^{-1}(r)$, $t = h(r)$, let q be the midpoint of pr and let s be the midpoint of rt. Then the composition hg is a hyperbolic isometry which translates along the line qs by the distance $2qs$.*

Proof See figure 3.1. Let p', r', t' be the respective feet of the perpendiculars from p, r, t to the line containing qs. Since $pq = qr = rs = st$ and since qrs is isosceles we see that $pp'q, rr'q, rr's, tt's$ are all congruent. Let the angles in these triangles be $\alpha, \beta, \pi/2$ as shown. Consider Fermi coordinates with respect to qs, based at r'. Then we have $p = (-a, b)$ and $t = (a, b)$ for some $a, b > 0$, where a is equal to the distance qs.

Consider a unit tangent vector (p, u) at p pointing towards p'. We consider the image of (p, u) under Dg, the derivative of g. Since g translates along pr, the vector

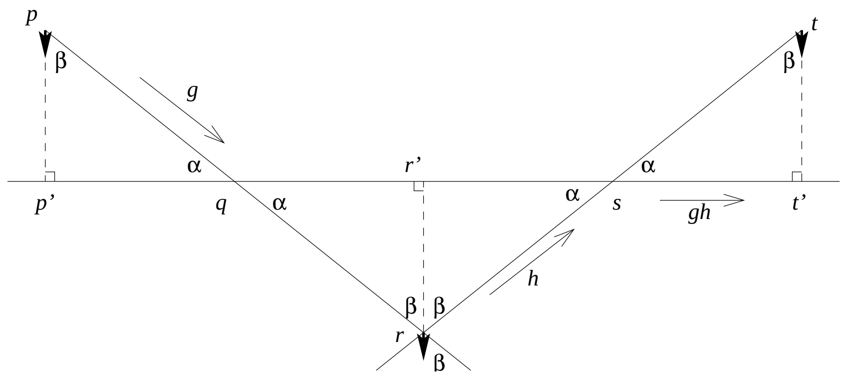

Figure 3.1: Situation of lemma 3.2.1.

$Dg(p, u)$ is based at r and points β clockwise of the direction of translation of g. It also points $\pi - \beta$ clockwise of the direction of translation of h. Thus $D(hg)(p, u)$ is based at t and points $\pi - \beta$ clockwise of the direction of translation of h. Thus (p, u) is based at $(-a, b)$ and points down towards p', while $D(hg)(p, u)$ is based at (a, b) and points down towards t'. But this is precisely the action of a translation along qs by distance $2a$. ∎

The main focus of our efforts in this section however is not on the composition of two but *four* isometries, namely in forming the *commutator* $[g, h] = ghg^{-1}h^{-1}$ of the isometries g and h. Note that, although g, h are only defined up to sign in $SL_2\mathbb{R}$, the commutator *is* a well-defined element of $SL_2\mathbb{R}$, and has a well-defined trace. (In fact we will see shortly that $[g, h]$ is well-defined in $\widetilde{PSL_2\mathbb{R}}$.)

Our next result characterises algebraically the geometric situation where g, h are hyperbolic and their axes intersect. A proof may be found in [30]: by computations after conjugating matrices to a simple standard form.

Lemma 3.2.2 *Let $g, h \in PSL_2\mathbb{R}$. The following are equivalent:*

(i) g, h are hyperbolic and their axes cross;

(ii) $\text{Tr}[g, h] < 2$. ∎

The next result describes the fixed points of a hyperbolic isometry. Here and throughout, for a hyperbolic isometry g, we denote its attractive and repulsive fixed points by a_g and r_g.

Lemma 3.2.3 *In the upper half plane,*

(i) if $g = [g_{ij}] \in SL_2\mathbb{R}$ represents a hyperbolic isometry, then

$$a_g + r_g = \frac{g_{11} - g_{22}}{g_{21}}, \qquad a_g r_g = -\frac{g_{12}}{g_{21}};$$

(ii) if $g = [g_{ij}] \in SL_2\mathbb{R}$ represents an elliptic isometry with fixed point p then

$$2\,\mathrm{Re}\,p = \frac{g_{11} - g_{22}}{g_{21}}, \qquad |p|^2 = -\frac{g_{12}}{g_{21}}.$$

Proof

(i) Clearly a_g, r_g are the solutions of

$$\frac{g_{11}z + g_{12}}{g_{21}z + g_{22}} = z,$$

so that

$$g_{21}z^2 + (g_{22} - g_{11})z - g_{12} = g_{21}(z - a_g)(z - r_g).$$

Expanding out and comparing coefficients of z gives the desired result.

(ii) The same proof applies once we note that the two solutions of the quadratic

$$\frac{g_{11}z + g_{12}}{g_{21}z + g_{22}} = z,$$

are p and $\bar{p}$. ∎

Next we describe the location of the axis of $[g, h]$ when it is hyperbolic and $\mathrm{Tr}[g, h] < -2$.

Lemma 3.2.4 *Suppose $g, h \in PSL_2\mathbb{R}$ and $\mathrm{Tr}[g, h] < -2$, so that g, h are hyperbolic and their axes intersect, and $[g, h]$ is also hyperbolic. Then $\mathrm{Axis}[g, h]$ does not intersect the axis of g or h. The fixed points of $[g, h]$ lie on the segment of the circle at infinity between a_g and a_h. The attractive fixed point $a_{[g,h]}$ is closer to a_g, and the repulsive fixed point $r_{[g,h]}$ is closer to a_h.*

The lemma states that the order of the fixed points on the circle at infinity is

$$a_h, r_{[g,h]}, a_{[g,h]}, a_g, r_h, r_g$$

up to cyclic permutation and reflection. See figure 3.2. Here and throughout we denote the translation distance of a hyperbolic isometry $\alpha \in PSL_2\mathbb{R}$ by d_α. And for a line l in $\mathbb{H}^2$, denote by R_l the reflection in l.

Proof We give the following elegant argument of Matelski in [46]; there is also a proof by computation. Let $p \in \mathbb{H}^2$ be the point of intersection of axes of g and h, and let $e \in PSL_2\mathbb{R}$ be a half-turn about p. Thus we have $ege = g^{-1}$ and $ehe = h^{-1}$. Consider he: this preserves $\mathrm{Axis}\,h$ but reverses its sense; it is therefore a half-turn

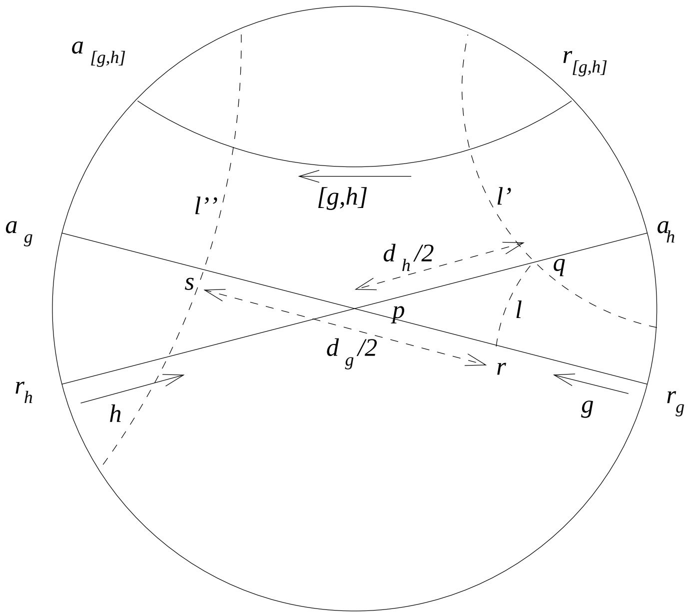

Figure 3.2: The location of $\mathrm{Axis}[g,h]$ in the case $\mathrm{Tr}[g,h] < -2$

about a point $q \in \mathrm{Axis}\, h$. In fact we see q lies on the same side of p as a_h, at a distance $d_h/2$ from p.

Now consider ghe. We have $(ghe)^2 = gh(ege)(ehe) = ghg^{-1}h^{-1} = [g,h]$, which is hyperbolic. Thus ghe is hyperbolic and has the same axis as $[g,h]$. So we only need show that the axis of ghe lies in the desired position.

Let l be the perpendicular from q to $\mathrm{Axis}\, g$, and r its foot. Let l' be the line through q perpendicular to l. Let s be a point along $\mathrm{Axis}\, g$ on the same side of r as a_g, and distance $d_g/2$ from r. Let l'' be the line through r perpendicular to $\mathrm{Axis}\, g$. So $R_{l'}R_l = he$; the composition of two reflections in two perpendicular lines meeting at q is a half-turn about q. And $R_l R_{l''} = g$; the composition of two reflections in lines perpendicular to $\mathrm{Axis}\, g$ being $d_g/2$ apart is a translation along $\mathrm{Axis}\, g$ by d_g. Thus $ghe = R_{l'}R_l^2 R_{l''} = R_{l'}R_{l''}$. So l' and l'' do not intersect, even at infinity (otherwise ghe would be elliptic or parabolic), and $\mathrm{Axis}\, ghe = \mathrm{Axis}[g,h]$ is the common perpendicular of l' and l''.

Now $l, l', \mathrm{Axis}[g,h], l'', \mathrm{Axis}\, g$ form a right-angled pentagon, as shown in figure 3.2. Thus $\mathrm{Axis}[g,h]$ must lie on the same side of $\mathrm{Axis}\, g$ as a_h, and on the same side of $\mathrm{Axis}\, h$ as a_g; and $[g,h]$ translates in the desired direction. ∎

Repeating the same argument when $[g,h]$ is parabolic, i.e. $a_{[g,h]} = r_{[g,h]}$, we see our right-angled pentagon now has a vertex at infinity and obtain:

Lemma 3.2.5 *Suppose $g, h \in PSL_2\mathbb{R}$ and $\mathrm{Tr}[g,h] = -2$, so that g, h are hyperbolic and their axes intersect, and $[g,h]$ is parabolic. Then fixed point at infinity of $[g,h]$*

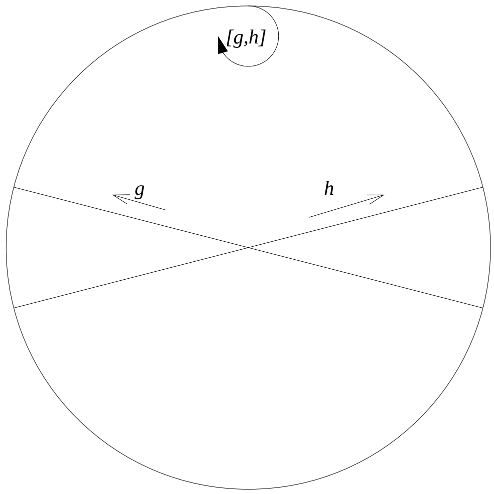

Figure 3.3: The location of Fix$[g,h]$ in the case Tr$[g,h] = -2$

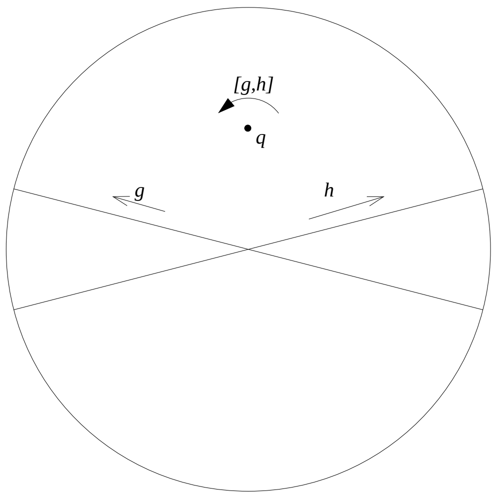

Figure 3.4: The location of Fix$[g,h]$ in the case Tr$[g,h] \in (-2,2)$

lies on the segment of the circle at infinity between a_g and a_h. The sense of the rotation is as shown in figure 3.3. ∎

Finally we consider the case when $-2 < \mathrm{Tr}[g,h] < 2$, i.e. g, h are hyperbolic, their axes cross, and $[g,h]$ is elliptic. The same construction shows that l', l'' intersect, and we obtain:

Lemma 3.2.6 *Suppose* $\mathrm{Tr}[g,h] \in (-2,2)$. *Then the fixed point q of $[g,h]$ lies in the region of $\mathbb{H}^2$ determined by* $\mathrm{Axis}(g), \mathrm{Axis}(h)$ *which is bounded by the arc on the circle at infinity between a_g and a_h.* ∎

3.3 $PSL_2\mathbb{R}$ and $\widetilde{PSL_2}\mathbb{R}$

We now make some general comments regarding these groups; see also [25] and [28]. Recall that $PSL_2\mathbb{R}$ may be identified with the unit tangent bundle $UT\mathbb{H}^2$, which is homeomorphic to $\mathbb{R}^2 \times S^1$, after choosing a basepoint $(y_0, u_0) \in UT\mathbb{H}^2$). We may therefore think of $PSL_2\mathbb{R}$ as the hyperbolic plane, with a circle of unit tangent vectors attached at each point. We have $\pi_1(PSL_2\mathbb{R}) \cong \mathbb{Z}$. Let p_1 be the projection map $PSL_2\mathbb{R} \longrightarrow \mathbb{H}^2$.

We may therefore consider the universal cover $\widetilde{PSL_2}\mathbb{R}$ with projection map $p_2 : \widetilde{PSL_2}\mathbb{R} \longrightarrow PSL_2\mathbb{R}$. We define an element $\tilde{x} \in \widetilde{PSL_2}\mathbb{R}$ as hyperbolic, elliptic or parabolic according to the type of $p_2(\tilde{x}) \in PSL_2\mathbb{R}$.

$$\left\{ \begin{array}{c} \widetilde{PSL_2}\mathbb{R} \\ \cong\updownarrow \\ \mathbb{H}^2 \times \mathbb{R}^1 \end{array} \right\} \xrightarrow{p_2} \left\{ \begin{array}{c} UT\mathbb{H}^2 \\ \cong\updownarrow \\ PSL_2\mathbb{R} \\ \cong\updownarrow \\ \mathbb{H}^2 \times S^1 \end{array} \right\} \xrightarrow{p_1} \mathbb{H}^2$$

We can think of this universal cover as the hyperbolic plane, with a line attached to each point, covering the circle of unit tangent vectors at that point. We can also think of elements of $\widetilde{PSL_2}\mathbb{R}$ as homotopy classes of paths in $UT\mathbb{H}^2$ starting at the basepoint. Since the basepoint is arbitrary, *every* path $c : [0,1] \longrightarrow UT\mathbb{H}^2$ (regardless of where it starts), i.e. every path in $\mathbb{H}^2$ with a unit tangent vector attached at each point, determines a unique element of $\widetilde{PSL_2}\mathbb{R}$, which we also denote c, hoping that not too much confusion will result from the abuse of notation. The projection of c to $PSL_2\mathbb{R}$ is the isometry sending the start tangent vector $c(0)$ to the end tangent vector $c(1)$.

An element α of $PSL_2\mathbb{R}$ has infinitely many lifts $\tilde{\alpha} \in \widetilde{PSL_2}\mathbb{R}$. These can all be taken to represent paths in $UT\mathbb{H}^2$ between the same start and end tangent vectors. However these paths will differ according to the number of times that the tangent vectors spin as the path in $UT\mathbb{H}^2$ is traversed.

The lifts of the identity $1 \in PSL_2\mathbb{R}$ form an infinite cyclic group, and correspond to those paths $c : [0,1] \longrightarrow UT\mathbb{H}^2$ which start and end at the same unit tangent vector. Any such path is clearly homotopic to a path $c : [0,1] \longrightarrow UT\mathbb{H}^2$ which projects to a constant path $p_1 \circ c : [0,1] \longrightarrow \mathbb{H}^2$. We see that c is then homotopic to the curve

$$c(t) = (x_0, e^{2\pi i n t} u_0)$$

for some $n \in \mathbb{Z}$. We denote the homotopy class of this curve $\mathbf{z}^n \in \widetilde{PSL_2}\mathbb{R}$. It is clear that the index notation agrees with the group structure on $\widetilde{PSL_2}\mathbb{R}$. We see

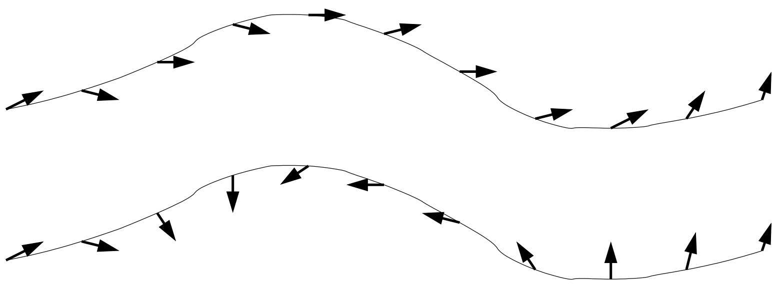

Figure 3.5: Paths in $UT\mathbb{H}^2 \cong PSL_2\mathbb{R}$, or equivalently, elements of $\widetilde{PSL_2}\mathbb{R}$. They project to the same element of $PSL_2\mathbb{R}$.

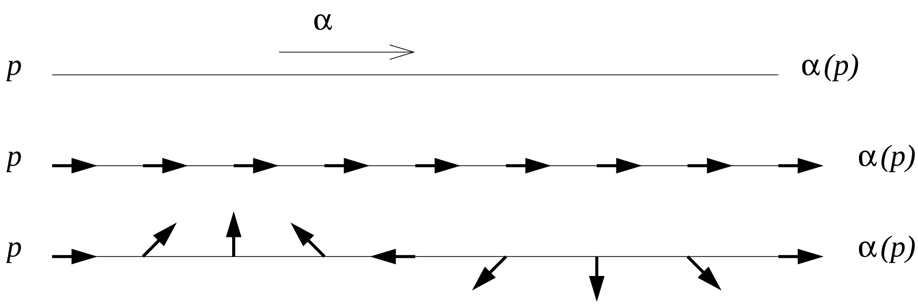

Figure 3.6: An isometry $\alpha \in PSL_2\mathbb{R}$; the simplest lift of α; a different lift of α.

that $\mathbf{z}$ commutes with every element of $\widetilde{PSL_2}\mathbb{R}$. In fact $\mathbf{z}$ generates the centre of $\widetilde{PSL_2}\mathbb{R}$, since $\pi_1(PSL_2\mathbb{R}) \cong \mathbb{Z}$ (see also [28]).

While every element has infinitely many lifts, there is a sense in which some lifts are simpler than others. For instance, the identity in $\widetilde{PSL_2}\mathbb{R}$ is, in some sense, the simplest lift of the identity in $PSL_2\mathbb{R}$.

If $\alpha \in PSL_2\mathbb{R}$ is hyperbolic then it translates by some distance d_α along an axis $\text{Axis}\,\alpha$ in $\mathbb{H}^2$. Let $c(t) \in PSL_2\mathbb{R}$ be the translation of (signed) hyperbolic distance td_α along $\text{Axis}\,\alpha$ in the same direction as α. Then $c : \mathbb{R} \longrightarrow PSL_2\mathbb{R}$ is a homomorphism with $c(1) = \alpha$, in fact the only homomorphism with this property. The path $c|_{[0,1]}$ in $PSL_2\mathbb{R}$ gives a unique element $\tilde{\alpha}$ of $\widetilde{PSL_2}\mathbb{R}$ which we take as our preferred or simplest lift. This lift can be thought of as a path of tangent vectors in $UT\mathbb{H}^2$, which travels along $\text{Axis}\,\alpha$ at speed d, with unit tangent vectors always pointing along $\text{Axis}\,\alpha$ in the direction of translation.

A similar set of ideas applies to parabolic isometries. If $\alpha \in PSL_2\mathbb{R}$ is parabolic then it translates along some horocycle h_α (here h_α is not unique). We may consider h_α with the induced Euclidean metric, so that α translates by Euclidean distance d. Again let $c(t) \in PSL_2\mathbb{R}$ be the parabolic isometry induced by translation along h_α by (signed) Euclidean distance td in the same direction as α. Again $c : \mathbb{R} \longrightarrow PSL_2\mathbb{R}$ is the unique homomorphism with $c(1) = \alpha$, and again $c|_{[0,1]}$ gives a unique element

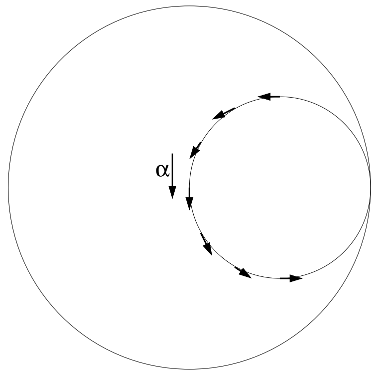

Figure 3.7: Simplest lift of a parabolic element of $PSL_2\mathbb{R}$

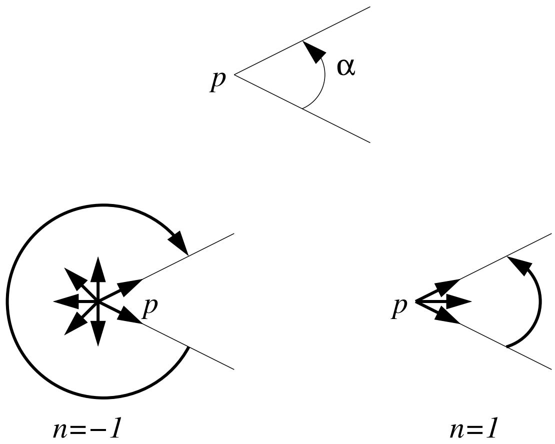

Figure 3.8: Simplest lifts of an elliptic element of $PSL_2\mathbb{R}$.

$\tilde{\alpha}$ of $\widetilde{PSL_2\mathbb{R}}$ which we take as our preferred lift. This lift $\tilde{\alpha}$ can be thought of as a path of tangent vectors travelling along h_α at speed d for time 1, with unit tangent vectors always pointing along h_α in the direction of translation.

However the situation for $\alpha \in PSL_2\mathbb{R}$ elliptic is quite different. If α is a rotation of angle $\theta \in (0, 2\pi)$ (measured anticlockwise) about a point $p \in \mathbb{H}^2$, then there are infinitely many homomorphisms $c : \mathbb{R} \longrightarrow PSL_2\mathbb{R}$ with $c(1) = \alpha$. For $n > 0$, let $c_n(t)$ be a rotation of angle $(\theta - 2\pi + 2n\pi)t$ (taken modulo 2π) about p. For $n < 0$, let $c_n(t)$ be a rotation of angle $(\theta - 2|n|\pi)t$ about p. This defines a different homomorphism for each $n \neq 0$, and the paths $\phi_n([0,1]), \phi_m([0,1])$ in $PSL_2\mathbb{R}$ are not homotopic (relative to endpoints) for $n \neq m$. Each $\phi_n|_{[0,1]}$ gives an element $\tilde{\alpha}_n \in \widetilde{PSL_2\mathbb{R}}$. The lift $\tilde{\alpha}_n$ can be thought of as a path of tangent vectors, based at p, and rotating with speed $\theta - 2\pi + 2n\pi$ (for $n > 0$) or $\theta - 2|n|\pi$ (for $n < 0$) for time 1. From this viewpoint there are two simplest lifts of α, namely $\tilde{\alpha}_1$ and $\tilde{\alpha}_{-1}$. The former is the simplest anticlockwise lift; the latter is the simplest clockwise lift.

We define the sets of simplest lifts of hyperbolic and parabolic elements of $PSL_2\mathbb{R}$

into $\widetilde{PSL_2\mathbb{R}}$ respectively as Hyp_0 and Par_0. For any hyperbolic $\tilde{\alpha} \in \widetilde{PSL_2\mathbb{R}}$, there is a unique $n \in \mathbb{Z}$ such that $\mathbf{z}^{-n}\tilde{x} \in \mathrm{Hyp}_0$. Thus we define $\mathrm{Hyp}_n = \mathbf{z}^n \,\mathrm{Hyp}_0$, and we see that the set of hyperbolic elements of $\widetilde{PSL_2\mathbb{R}}$ is the disjoint union of the sets Hyp_n. We consider an element $\tilde{\alpha}$ of Hyp_n as a translation of length d along $\mathrm{Axis}\,\alpha$ with an added twist of $2n\pi$. Similarly we define $\mathrm{Par}_n = \mathbf{z}^n \,\mathrm{Par}_0$, and we consider an element $\tilde{\alpha}$ of Par_n as a translation along a horocycle h_α with an added twist of $2n\pi$. We may further distinguish between Par_n^+ and Par_n^-, the rotations about points at infinity whose projections to $PSL_2\mathbb{R}$ are anticlockwise and clockwise respectively.

We make some similar definitions for elliptic elements. Let the set of simplest anticlockwise lifts of elliptic elements be Ell_1. The set of simplest clockwise lifts of elliptic elements is denoted Ell_{-1}. For $n > 0$ we define $\mathrm{Ell}_n = \mathbf{z}^{n-1}\,\mathrm{Ell}_1$ and $\mathrm{Ell}_{-n} = \mathbf{z}^{-n+1}\,\mathrm{Ell}_{-1}$. Note Ell_0 is not defined, and actually $\mathbf{z}\,\mathrm{Ell}_{-1} = \mathrm{Ell}_1$. The Ell_n are disjoint and contain all elliptic elements of $\widetilde{PSL_2\mathbb{R}}$. For $n > 0$ (resp. $n < 0$), Ell_n consists of all rotations through angles between $2\pi(n-1)$ and $2\pi n$ (resp. between $2\pi n$ and $2\pi(n+1)$).

We may consider the exponential map $\exp : \mathfrak{psl}_2\mathbb{R} \longrightarrow PSL_2\mathbb{R}$ and its lift $\widetilde{\exp} : \widetilde{\mathfrak{psl}_2\mathbb{R}} \longrightarrow \widetilde{PSL_2\mathbb{R}}$. Then we see that

$$\widetilde{\exp}\left(\widetilde{\mathfrak{psl}_2\mathbb{R}}\right) = \bigcup_{n\in\mathbb{Z}} \{\mathbf{z}^n\} \cup \mathrm{Hyp}_0 \cup \mathrm{Par}_0 \cup \bigcup_{n\in\mathbb{Z}\setminus\{0\}} \mathrm{Ell}_n \,.$$

Considering that the hyperbolic and elliptic elements form two disjoint 3-dimensional subspaces of the 3-dimensional space $PSL_2\mathbb{R}$, and that their common 2-dimensional boundary is the space of parabolic elements, we may draw a schematic diagram of the universal cover $\widetilde{PSL_2\mathbb{R}}$ as in figure 3.9.

We conclude this introduction to $\widetilde{PSL_2\mathbb{R}}$ with a proposition about commutators which will be crucial in the sequel. Recall $[\alpha, \beta] = \alpha\beta\alpha^{-1}\beta^{-1}$.

Lemma 3.3.1 *Let $\alpha, \beta \in PSL_2\mathbb{R}$. Then $[\alpha, \beta]$ has a well-defined lift to $\widetilde{PSL_2\mathbb{R}}$. That is, any two sets of lifts $\tilde{\alpha}_1, \tilde{\beta}_1$ and $\tilde{\alpha}_2, \tilde{\beta}_2$ satisfy $[\tilde{\alpha}_1, \tilde{\beta}_1] = [\tilde{\alpha}_2, \tilde{\beta}_2]$.*

Proof Let $\tilde{\alpha}_2 = \mathbf{z}^a\tilde{\alpha}_1$, $\tilde{\beta}_2 = \mathbf{z}^b\tilde{\beta}_1$. Since $\mathbf{z}$ commutes with every element of $\widetilde{PSL_2\mathbb{R}}$ we see that

$$\begin{aligned}[\tilde{\alpha}_2, \tilde{\beta}_2] &= \tilde{\alpha}_2\tilde{\beta}_2\tilde{\alpha}_2^{-1}\tilde{\beta}_2^{-1} = \mathbf{z}^a\tilde{\alpha}_1\mathbf{z}^b\tilde{\beta}_1\tilde{\alpha}_1^{-1}\mathbf{z}^{-a}\tilde{\beta}_1^{-1}\mathbf{z}^{-b} \\ &= \tilde{\alpha}_1\tilde{\beta}_1\tilde{\alpha}_1^{-1}\tilde{\beta}_1^{-1} = [\tilde{\alpha}_1, \tilde{\beta}_1]\end{aligned}$$

as desired. ∎

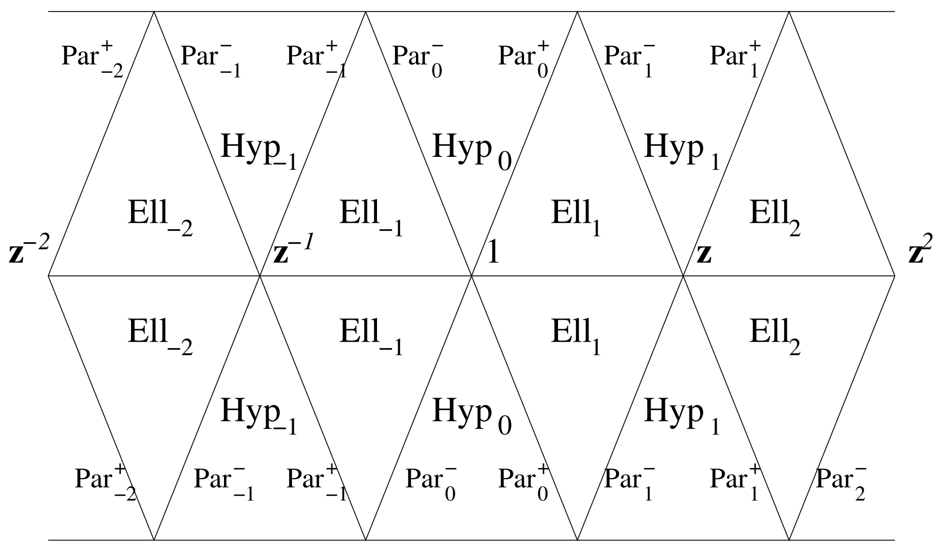

Figure 3.9: Schematic diagram of $\widetilde{PSL_2\mathbb{R}}$; think of this as a solid tube. The centre line represents rotations about the basepoint.

3.4 Derivatives of Isometries of $\mathbb{H}^2$

We now consider the action of isometries of $\mathbb{H}^2$ in more detail. In the proof of lemma 3.2.1 we considered derivatives of isometries, and their effect on unit tangent vectors; similar techniques will be used later.

An isometry $\alpha \in PSL_2\mathbb{R}$ has a derivative $D\alpha$ which can be considered as a map $UT\mathbb{H}^2 \longrightarrow UT\mathbb{H}^2$. We will consider the "twist" involved in this action. In fact we will define more generally the twist of $\tilde{\alpha} \in \widetilde{PSL_2\mathbb{R}}$ at $y \in \mathbb{H}^2$.

First we define the *twist of a vector field $\mathcal{V}$ along the curve c*. Consider a smooth curve $c : [0,1] \longrightarrow \mathbb{H}^2$ and a smooth unit tangent vector field $\mathcal{V} : [0,1] \longrightarrow UT\mathbb{H}^2$ with $p_1 \circ \mathcal{V} = c$ (recall p_1 is the projection $UT\mathbb{H}^2 \longrightarrow \mathbb{H}^2$). There is a well-defined velocity vector $\frac{dc}{dt} \in T\mathbb{H}^2$ at each point, and we may rescale this to be a unit vector. This gives a unit tangent vector field $\hat{c} : [0,1] \longrightarrow UT\mathbb{H}^2$ on the curve c, where $p_1 \circ \hat{c} = c$. Intuitively $\hat{c}(t)$ tells us in which direction we are travelling at time t. We may consider the angle $\theta(t)$ from $\hat{c}(t)$ to $\mathcal{V}(t)$, measured anticlockwise, at the time t. We have many choices for $\theta(0)$ (all choices differ by multiples of 2π), but making an arbitrary choice for $\theta(0)$ and requiring θ to be continuous determines θ completely. We then define $\theta(1) - \theta(0)$ to be the *twist of the vector field $\mathcal{V}$ along the curve c*. It's clear this is independent of the particular choice of $\theta(0)$.

Now given y in $\mathbb{H}^2$ and $\tilde{\alpha} \in \widetilde{PSL_2\mathbb{R}}$ we define the *twist of $\tilde{\alpha}$ at y*, denoted $\mathrm{Tw}(\tilde{\alpha}, y)$, as follows. Let $\tilde{\alpha}$ project to $\alpha \in PSL_2\mathbb{R}$ (under $p_2 : \widetilde{PSL_2\mathbb{R}} \longrightarrow PSL_2\mathbb{R}$). Let $c : [0,1] \longrightarrow \mathbb{H}^2$ be a constant speed geodesic in the hyperbolic plane between y and $\alpha(y)$ (the speed could be 0). There is a vector field $[0,1] \longrightarrow UT\mathbb{H}^2$ which lies

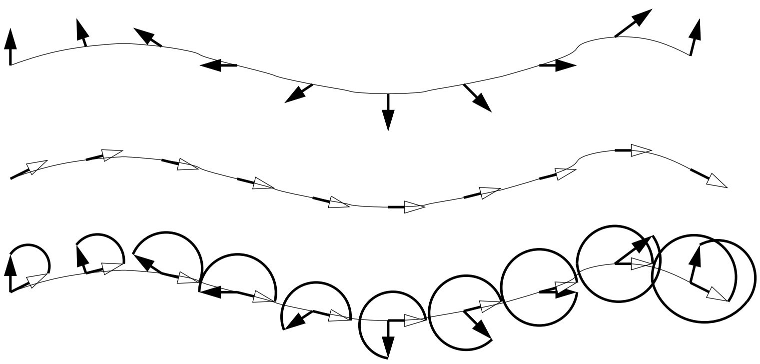

Figure 3.10: The twist of a vector field along a curve.

in the homotopy class of $\tilde{\alpha}$ and projects to c (under $p_1 : UT\mathbb{H}^2 \to \mathbb{H}^2$); let this be $\mathcal{V}$. Then $\text{Tw}(\tilde{\alpha}, y)$ is defined to be the twist of the vector field $\mathcal{V}$ along the geodesic c.

Intuitively, $\text{Tw}(\tilde{\alpha}, y)$ describes how the tangent vector at y is moved by $\tilde{\alpha}$, compared to parallel translation along the geodesic from y to $\tilde{\alpha}(y)$. It is clear that the definition of $\text{Tw}(\tilde{\alpha}, y)$ depends only on $\tilde{\alpha}$ and y, and not on the choice of the particular vector field $\mathcal{V}$. Now for $\alpha \in PSL_2\mathbb{R}$ we may define $\text{Tw}(\alpha, y)$ to be equal to $\text{Tw}(\tilde{\alpha}, y)$ for any lift $\tilde{\alpha}$ of α, with the angle taken modulo 2π.

We will now consider the twist of various isometries. Throughout this section, let L_θ denote the locus of points y such that $\text{Tw}(\tilde{\alpha}, y) = \theta$. We will determine the loci L_θ.

First consider a hyperbolic isometry α. In terms of Fermi coordinates with respect to l, a hyperbolic isometry α with axis l acts as $(x, h) \mapsto (x + d_\alpha, h)$, and hence preserves the curves at constant height h from l. Recall that curves of constant distance from l appear in the upper half plane model as Euclidean circles or lines. If l is the imaginary axis in the upper half plane model, then the curves of constant curvature parallel to l are Euclidean lines through the point 0. Curves of constant distance have constant curvature, and share the same endpoints at infinity as l. They are parameterised by their signed height h from l and foliate $\mathbb{H}^2$. Denote the constant distance curve at height h from l by $C_h(l)$.

Using the distance formula for Fermi coordinates (see equation 3.1, see further e.g. [10]), we see that the distance between (x, h) and $(x + d_\alpha, h)$ is given by

$$\cosh^{-1}\left(\cosh^2 h\,(\cosh d_\alpha - 1) + 1\right).$$

So α translates by distance d_α along the axis l, but elsewhere moves a point by *more* than d_α. However along the curves $C_h(l)$ there is a twisting effect involved as well. Consider $y \in \mathbb{H}^2$ at height h from l. The geodesic segment γ between y and $\alpha(y)$ will

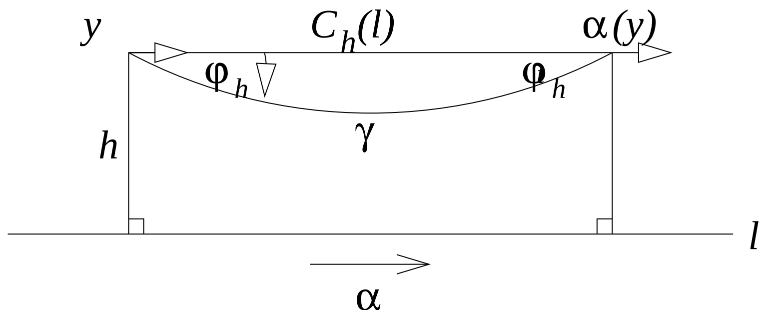

Figure 3.11: Twist of a hyperbolic isometry α at $y \in \mathbb{H}^2$ not on $l = \text{Axis}(\alpha)$.

be strictly shorter than the constant distance curve $C_h(l)$ between the same points. In fact because of the negative curvature of $\mathbb{H}^2$ the geodesic segment γ lies on the same side of $C_h(l)$ as the line l. Let ϕ_h denote the signed angle in $(-\pi/2, \pi/2)$ from $C_h(l)$ to γ at y. See figure 3.11.

Consider the derivative of α acting on the unit tangent bundle of the upper half plane. A unit tangent vector at y, pointing along $C_h(l)$, is taken to a unit tangent vector at $\alpha(y)$ also pointing along $C_h(l)$ in the same direction. Relative to the geodesic γ, there is a twist of $2\phi_h$, thus $\text{Tw}(y, \alpha) = 2\phi_h$. It is clear that the angle ϕ_h is strictly decreasing as h increases. Taking signed distance appropriately, $\phi_h \to \mp\frac{\pi}{2}$ as $h \to \pm\infty$. So the locus L_θ is precisely the curve at the height h such that $\theta = 2\phi_h$.

If we consider the simplest lift $\tilde{\alpha} \in \text{Hyp}_0$ of α, then $\tilde{\alpha}$ can simply be thought of as a path of tangent vectors always pointing along $C_h(l)$. Thus $\text{Tw}(y, \alpha) = 2\phi_h$. We record this observation. Note that for $\tilde{\alpha} \in \text{Hyp}_n$ generally we must adjust by an appropriate multiple of 2π.

Lemma 3.4.1 *Let $\tilde{\alpha} \in \text{Hyp}_0$ be the simplest lift of a hyperbolic isometry and let θ be a real number. Then L_θ is a curve of constant distance from the axis of α, for each $-\pi < \theta < \pi$. For $\theta \leq -\pi$ or $\theta \geq \pi$, the locus L_θ is empty. The curves L_θ are disjoint and foliate $\mathbb{H}^2$.* ■

Now consider an elliptic isometry $\tilde{\alpha} \in \text{Ell}_1$ rotating by an angle ψ $(0 < \psi < 2\pi)$ about a point q. First suppose $\psi \in (0, \pi)$. Then for the projection $\alpha \in PSL_2\mathbb{R}$ of $\tilde{\alpha}$ we have $\text{Tr}\,\alpha = \pm 2\cos(\psi/2)$ and α preserves any hyperbolic circle $C_h(q)$ of radius h about q. The circles $C_h(q)$, which also have constant curvature, will be analogous to the constant distance curves in the hyperbolic case. The length of the curve $C_h(q)$ between y and $\alpha(y)$ is strictly longer than the length of the geodesic segment γ between the same two points. Again consider the action of the derivative $D\alpha$ on $UT\mathbb{H}^2$. A unit vector at y pointing along $C_h(q)$ is clearly taken by $D\alpha$ to a

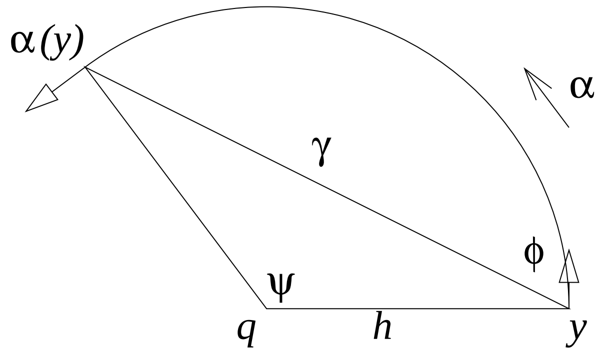

Figure 3.12: Twist of an elliptic isometry.

unit vector at $\alpha(y)$ also pointing along $C_h(q)$ in the same direction. A unit vector at y pointing along γ is twisted by an angle $2\phi_h$ $(0 < 2\phi < 2\pi)$, where ϕ_h is the angle from $C_h(q)$ to γ at y. As $h \to 0$ we have $\phi_0 = \psi/2$ (a Euclidean picture), and ϕ_h is strictly monotone increasing to $\pi/2$ as h increases, so we have a similar result. The twist $\mathrm{Tw}(\alpha, y)$ depends only on h, and increases from ψ to π as h increases.

If $\psi = \pi$ then $\tilde{\alpha} \in \mathrm{Ell}_1$ is a half turn about q. So ψ preserves all the lines through q. A unit vector at a point y on such a line, pointing along the line away from q, is taken by $D\alpha$ to a unit vector at $\alpha(y)$, also pointing away from q, and is twisted by π. So at every $y \in \mathbb{H}^2$, $\mathrm{Tw}(\tilde{\alpha}, y) = \pi$.

And if $\tilde{\alpha} \in \mathrm{Ell}_1$ rotates by angle $\psi \in (\pi, 2\pi)$, then consider $\tilde{\alpha}' = \mathbf{z}^{-1}\tilde{\alpha} \in \mathrm{Ell}_{-1}$, which rotates by angle $\psi' = \psi - 2\pi \in (-\pi, 0)$. So we revert to the first case, with opposite orientation, noting $\mathrm{Tw}(\tilde{\alpha}', y) = \mathrm{Tw}(\tilde{\alpha}, y) - 2\pi$. So for $-\pi < \theta < \psi - 2\pi$, L_θ is a circle centred at q; for $\theta = \psi - 2\pi$ it is q; and otherwise L_θ is empty. So we obtain the following result.

Lemma 3.4.2 *Let $\tilde{\alpha} \in \mathrm{Ell}_1$ be a rotation of angle $\psi \in (0, 2\pi)$ about $q \in \mathbb{H}^2$.*

(i) If $\psi \in (0, \pi)$ then the locus L_θ is

- *(a) a hyperbolic circle centred at q, for $\psi < \theta < \pi$,*
- *(b) the point q, for $\theta = \psi$, and*
- *(c) empty otherwise.*

(ii) If $\psi = \pi$, then $L_\pi = \mathbb{H}^2$ and for every $\theta \neq \pi$, L_θ is empty.

(iii) If $\psi \in (\pi, 2\pi)$ then the locus L_θ is

- *(a) a hyperbolic circle centred at q, for $\pi < \theta < \psi$,*
- *(b) the point q, for $\theta = \psi$, and*

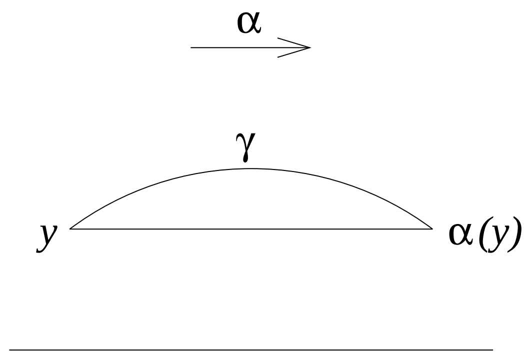

Figure 3.13: Twist of a parabolic isometry.

(c) empty otherwise. ■

Finally, consider a parabolic isometry $\tilde{\alpha} \in \mathrm{Par}_0^+$, which rotates anticlockwise about some point q at infinity. Extending our previous notation, we similarly obtain a foliation of $\mathbb{H}^2$ by horocycles $C_h(q)$ about q, parametrised by some arbitrary height function h; if we normalise so $q = \infty$ in the upper half plane model, we may take h simply to be the y-coordinate. Once more the projection α of $\tilde{\alpha}$ preserves the horocycles $C_h(q)$ and, for $y \in C_h(q)$, the curve along $C_h(q)$ between y and $\alpha(y)$ is longer than the geodesic segment γ connecting the same two points. Again the twist $\mathrm{Tw}(\tilde{\alpha}, y)$ is given by $2\phi_h$ where ϕ_h is the angle from $C_h(q)$ to γ at y. The angle ϕ_h is monotone decreasing from $\pi/2$ to 0 as h increases from 0 to ∞. If $\tilde{\alpha} \in \mathrm{Par}_0^-$, i.e. $\tilde{\alpha}$ rotates clockwise, then ϕ_h is monotone increasing from $-\pi/2$ to 0 as h increases. Multiplication by $\mathbf{z}^n$ has the expected effect.

Lemma 3.4.3 *Let $\tilde{\alpha} \in \mathrm{Par}_0^+$ be an anticlockwise rotation about q at infinity. Then locus L_θ is a horocycle about q, for each $0 < \theta < \pi$. Otherwise L_θ is empty. The curves L_θ are disjoint and foliate $\mathbb{H}^2$.*

If $\tilde{\alpha} \in \mathrm{Par}_0^-$ then L_θ is also a horocycle, for each $-\pi < \theta < 0$. ■

We summarise with the following proposition.

Proposition 3.4.4

$$\mathrm{Tw}(\mathrm{Hyp}_n, \mathbb{H}^2) = ((2n-1)\,\pi, (2n+1)\,\pi)$$
$$\mathrm{Tw}(\mathrm{Par}_n, \mathbb{H}^2) = ((2n-1)\,\pi, (2n+1)\,\pi)$$
$$\mathrm{Tw}(\mathrm{Ell}_n, \mathbb{H}^2) = \begin{cases} \Big((2n-2)\pi, 2n\pi\Big) & \text{for } n > 0 \\ \Big(-2|n|\pi, (-2|n|+1)\pi\Big) & \text{for } n < 0 \end{cases}$$

These are the values of Tw *over the subsets* Hyp_n, Par_n, Ell_n *of* $\widetilde{PSL_2\mathbb{R}}$ *and over all points in* $\mathbb{H}^2$. ■

3.5 Milnor's angle function

In his 1957 paper [49] Milnor introduced a concept of the "angle" associated to an element of $GL_2^+\mathbb{R}$. This concept applies to matrices in $SL_2\mathbb{R}$. It is related to the notion of twisting discussed above, but entirely algebraic. We summarise his results: see his paper for details.

A matrix $\alpha \in SL_2\mathbb{R}$ can be written uniquely in the form

$$\alpha = R(\alpha)S(\alpha)$$

where $R, S \in SL_2\mathbb{R}$, R is orthogonal and S is symmetric positive definite. Since R is orthogonal, it is of the form

$$R(\alpha) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

for some θ. This θ can be thought of as the angle of rotation of α. If we project R into $PSL_2\mathbb{R}$, then it acts on the upper half plane model of $\mathbb{H}^2$ as a rotation of angle 2θ about i: this follows from $R(\alpha)(i) = i$ and $\operatorname{Tr} R(\alpha) = 2\cos\theta$. So the map $R : SL_2\mathbb{R} \longrightarrow SO_2\mathbb{R}$ is a retraction, from which we extract an angle θ (taken modulo 2π); this lifts to a retraction $\tilde{R} : \widetilde{SL_2\mathbb{R}} \longrightarrow \widetilde{SO_2\mathbb{R}}$, from which we wish to extract an angle (as a real number, not modulo 2π). This angle is just coming from the exponential map. Since $SO_2\mathbb{R} \cong S^1$, we have $\widetilde{SO_2\mathbb{R}} \cong \mathbb{R}$ and $\mathfrak{so}_2\mathbb{R} \cong \mathbb{R}$, so the exponential map

$$\exp : \mathbb{R} \longrightarrow SO_2\mathbb{R}, \qquad \theta \mapsto \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

lifts to $\widetilde{\exp} : \mathbb{R} \longrightarrow \widetilde{SO_2\mathbb{R}} \subset \widetilde{PSL_2\mathbb{R}}$. This map $\widetilde{\exp}$ takes θ to a rotation of angle 2θ about i. We define the angle function $\Theta : \widetilde{PSL_2\mathbb{R}} \longrightarrow \mathbb{R}$ by

$$\Theta(\tilde{\alpha}) = \widetilde{\exp}^{-1}\left(\tilde{R}(\alpha)\right).$$

From $R(\alpha^{-1}) = R(\alpha)^{-1}$ and continuity we can deduce $\tilde{R}(\tilde{\alpha}^{-1}) = \tilde{R}(\tilde{\alpha})^{-1}$, so that $\Theta(\tilde{\alpha}^{-1}) = -\Theta(\tilde{\alpha})$. Multiplication by $\mathbf{z}$ just adds π to the angle, so $\Theta(\mathbf{z}^n\tilde{\alpha}) = n\pi + \Theta(\tilde{\alpha})$.

Milnor proved that Θ is approximately additive:

Theorem 3.5.1 (Milnor [49]) *Take $\tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_k \in \widetilde{PSL_2\mathbb{R}}$. Then the function Θ satisfies*

$$\left|\Theta\left(\tilde{\alpha}_1\tilde{\alpha}_2\cdots\tilde{\alpha}_k\right) - \Theta\left(\tilde{\alpha}_1\right) - \Theta\left(\tilde{\alpha}_2\right) - \cdots - \Theta\left(\tilde{\alpha}_k\right)\right| < (k-1)\frac{\pi}{2}.$$

■

For the rest of this section, we use this approximate additivity to give bounds on Θ in certain useful cases.

Lemma 3.5.2 *Let k, n be non-negative integers satisfying $2-2k-n<0$. Suppose*

$$g_1, h_1, g_2, h_2, \dots, g_k, h_k, c_1, \dots, c_n \in PSL_2\mathbb{R}$$

satisfy

$$[g_1,h_1][g_2,h_2]\cdots[g_k,h_k]c_1c_2\cdots c_n = 1.$$

Let $\tilde{g}_i, \tilde{h}_i \in \widetilde{PSL_2\mathbb{R}}$ be arbitrary lifts of g_i, h_i and let $\tilde{c}_i$ be a lift of c_i satisfying $|\Theta(\tilde{c}_i)| \le \frac{\pi}{2}$. Then

$$\left[\tilde{g}_1, \tilde{h}_1\right]\left[\tilde{g}_2, \tilde{h}_1\right]\cdots\left[\tilde{g}_k, \tilde{h}_k\right]\tilde{c}_1\tilde{c}_2\cdots\tilde{c}_n = \mathbf{z}^m,$$

where $|m| \le -2+2k+n$.

Proof Clearly the product $[\tilde{g}_1,\tilde{h}_1]\cdots[\tilde{g}_k,\tilde{h}_k]\tilde{c}_1\cdots\tilde{c}_n$ is a lift of $1 \in PSL_2\mathbb{R}$, hence is equal to $\mathbf{z}^m$ for some $m \in \mathbb{Z}$.

Now both

$$\Theta\left(\left[\tilde{g}_1, \tilde{h}_1\right]\left[\tilde{g}_2, \tilde{h}_2\right]\cdots\left[\tilde{g}_k, \tilde{h}_k\right]\tilde{c}_1\tilde{c}_2\cdots\tilde{c}_n\right)$$

and

$$\Theta\left(\left[\tilde{g}_1, \tilde{h}_1\right]\left[\tilde{g}_2, \tilde{h}_2\right]\cdots\left[\tilde{g}_k, \tilde{h}_k\right]\tilde{c}_1\tilde{c}_2\cdots\tilde{c}_{n-1}\right)+\Theta\left(\tilde{c}_n\right)$$

must be integer multiples of π. For the first expression this is obvious, since the group elements in brackets multiply to a power of $\mathbf{z}$. For the second expression, note that the two group elements in brackets in the second expression are lifts of inverses, hence are inverses in $PSL_2\mathbb{R}$, up to multiplication by a power of $\mathbf{z}$, and that $\Theta(\tilde{\alpha}^{-1}) = -\Theta(\tilde{\alpha})$.

By approximate additivity, these two quantities may differ by at most $\pi/2$, hence they are equal. Then by approximate additivity applied $4k+n-2$ times, (note that the inverses in the commutators cancel out), we have

$$\begin{aligned}
\pi(-2k-n+1) &\le \Theta(\tilde{c_1}) + \cdots + \Theta(\tilde{c_{n-1}}) + (-4k-n+2)\frac{\pi}{2} + \Theta(\tilde{c_n}) \\
&< \Theta\left(\left[\tilde{g_1}, \tilde{h_1}\right]\left[\tilde{g_2}, \tilde{h_1}\right]\cdots\left[\tilde{g_k}, \tilde{h_k}\right]\tilde{c_1}\tilde{c_2}\cdots\tilde{c_{n-1}}\right) + \Theta\left(\tilde{c_n}\right) \\
&< \Theta(\tilde{c_1}) + \cdots + \Theta(\tilde{c_{n-1}}) + (4k+n-2)\frac{\pi}{2} + \Theta(\tilde{c_n}) \\
&\le \pi\left(2k+n-1\right).
\end{aligned}$$

Since this sum must be an integer multiple of π, we have

$$\Theta\left(\left[\tilde{g_1}, \tilde{h_1}\right]\left[\tilde{g_2}, \tilde{h_1}\right]\cdots\left[\tilde{g_k}, \tilde{h_k}\right]\tilde{c_1}\tilde{c_2}\cdots\tilde{c_n}\right) = (-2k-n+2)\pi, \dots, \text{ or } (2k+n-2)\pi$$

as required. ∎

Note that the above argument does not work for $2k+n \leq 2$. For instance, with $k=0, n=2$, setting $\tilde{c}_1 = \tilde{c}_2 = \widetilde{\exp}(\pi/2)$ provides a counterexample. Then both $\tilde{c}_1, \tilde{c}_2$ are half-turns about i in the upper half plane. We have $c_1 c_2 = 1 \in PSL_2\mathbb{R}$ but $\tilde{c}_1 \tilde{c}_2 = \mathbf{z}$.

Corollary 3.5.3 *If $g, h \in PSL_2\mathbb{R}$ then (noting $[g,h]$ is well-defined in $\widetilde{PSL_2\mathbb{R}}$)*

$$\Theta\left([g,h]\right) \in \left[-\frac{3\pi}{2}, \frac{3\pi}{2}\right].$$

PROOF Take $\tilde{c} \in \widetilde{PSL_2\mathbb{R}}$ projecting to $c \in PSL_2\mathbb{R}$ such that $[g,h]c = 1 \in PSL_2\mathbb{R}$ and such that $-\frac{\pi}{2} \leq \Theta(\tilde{c}) \leq \frac{\pi}{2}$. As in the previous proof, both $\Theta([g,h]\tilde{c})$ and $\Theta([g,h]) + \Theta(\tilde{c})$ are integer multiples of π; but by approximate additivity they differ by at most $\pi/2$; hence they are equal. By the previous result $\Theta([g,h]c) = -\pi, 0$ or π. By approximate additivity then

$$\left|\Theta\left([g,h]\right)\right| = \left|\Theta\left([g,h]\tilde{c}\right) - \Theta\left(\tilde{c}\right)\right| \leq \left|\Theta\left([g,h]\tilde{c}\right)\right| + \left|\Theta\left(\tilde{c}\right)\right| \leq \frac{3\pi}{2}$$

as desired. ∎

Note actually it will follow from proposition 3.7.2 that $\Theta([g,h]) \in (-3\pi/2, 3\pi/2)$.

3.6 A geometric interpretation for Θ

We now show a direct correspondence between Milnor's algebraic function Θ, and our geometric notion of twisting. Although it seems that similar ideas have been used previously, for instance in [28], [62], it seems this has not been described explicitly before.

Lemma 3.6.1 *In the upper half plane model, the geodesic with endpoints at infinity a, b passes through i if and only if $ab = -1$.*

PROOF The geodesic is a Euclidean semicircle, and hence passes through i if and only if a, b, i form a right-angled triangle with right angle at i. By Pythagoras' theorem (see figure 3.14) this is equivalent to

$$a^2 + 1 + b^2 + 1 = (a-b)^2$$

which simplifies to $ab = -1$. ∎

Lemma 3.6.2 *An isometry of $\mathbb{H}^2$ is represented by a symmetric positive definite matrix in $SL_2\mathbb{R}$ other than the identity if and only if it is hyperbolic and its axis passes through i.*

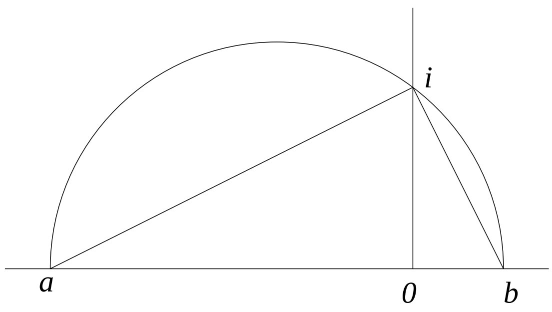

Figure 3.14: A geodesic passes through i if and only if $ab = -1$.

PROOF Take a symmetric positive definite matrix $\alpha = \begin{bmatrix} a & b \\ b & d \end{bmatrix} \neq 1 \in SL_2\mathbb{R}$. First we claim α is hyperbolic. We have $ad - b^2 = 1$, so $ad = 1 + b^2 \geq 1$. Since α and $-\alpha$ project to the same element of $PSL_2\mathbb{R}$, we may assume $a, d > 0$. Then we have $a + d \geq 2\sqrt{ad} \geq 2$. If equality holds then $a = d$ and $ad = 1$, so $a = d = 1$; then $ad - b^2 = 1$ implies $b = 0$, so S is the identity, contrary to assumption. So $\operatorname{Tr} \alpha > 2$ and α is hyperbolic.

Now by lemma 3.2.3, the fixed points of α multiply to $-b/b = -1$. The axis of α is the geodesic connecting these two fixed points, which by the previous lemma passes through i.

Conversely, suppose a matrix $\alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2\mathbb{R}$ represents a hyperbolic isometry with axis passing through i, hence with fixed points multiplying to -1. By lemma 3.2.3 again, the fixed points multiply to $-b/c = -1$. Thus $b = c$ and α is symmetric. Since α and $-\alpha$ both lie in $SL_2\mathbb{R}$ and represent the same isometry, we may choose the matrix with positive trace. That matrix has positive trace and determinant, hence both eigenvalues are positive, so it is positive definite. ■

We can now state the geometric interpretation of Θ.

Proposition 3.6.3 *Let $\tilde{\alpha} \in \widetilde{PSL_2\mathbb{R}}$ which projects to $\alpha \in PSL_2\mathbb{R}$. Then*

$$\Theta(\tilde{\alpha}) = \frac{1}{2} \operatorname{Tw}\left(\tilde{\alpha}, \alpha^{-1}(i)\right).$$

PROOF From the definition of Θ we have

$$\tilde{\alpha} = \widetilde{\exp}\left(\Theta(\tilde{\alpha})\right) \tilde{S}(\tilde{\alpha})$$

where $\tilde{S}(\tilde{\alpha}) \in \operatorname{Hyp}_0 \cup \{1\}$ is a lift of a symmetric positive definite matrix $S(\tilde{\alpha})$.

Consider now the action of $\tilde{\alpha}$ on the hyperbolic plane. Since $S(\tilde{\alpha})$ is symmetric positive definite, by 3.6.2 it is either the identity, or a translation along an axis passing through i. Then $\tilde{S}(\tilde{\alpha}) \in \mathrm{Hyp}_0 \cup \{1\}$ is the simplest lift of this translation. This action is followed by that of $\widetilde{\exp}(\Theta(\tilde{\alpha}))$, which is a rotation of angle $2\Theta(\tilde{\alpha})$ about i. So the overall action of $\tilde{\alpha}$ is to translate from $\alpha^{-1}(i)$ to i, and then rotate by angle $2\Theta(\tilde{\alpha})$.

Let $c : [0,1] \longrightarrow \mathbb{H}^2$ be a constant speed parameterisation of the geodesic segment from $\alpha^{-1}(i)$ to i. Let $\gamma : [0,1] \longrightarrow UT\mathbb{H}^2$ by a path which projects to c, i.e. $p_1 \circ \gamma = c$, and which has twist $2\Theta(\tilde{\alpha})$ along c. Then it is clear that γ represents the homotopy class of $\tilde{\alpha}$, so the twist of $\tilde{\alpha}$ at $\alpha^{-1}(i)$ is equal to the twist of γ along c, which is $2\Theta(\tilde{\alpha})$. ∎

Corollary 3.6.4 *Over the subsets* Hyp_n, Par_n, Ell_n *of* $\widetilde{PSL_2\mathbb{R}}$, *the function* Θ *takes the following values:*

$$\Theta(\mathrm{Hyp}_n) = \left(\left(n - \frac{1}{2}\right)\pi, \left(n + \frac{1}{2}\right)\pi\right)$$
$$\Theta(\mathrm{Par}_n) = \left(\left(n - \frac{1}{2}\right)\pi, \left(n + \frac{1}{2}\right)\pi\right)$$
$$\Theta(\mathrm{Ell}_n) = \begin{cases} \big((n-1)\pi, n\pi\big) & \text{for } n > 0 \\ \big(-|n|\pi, (-|n|+1)\pi\big) & \text{for } n < 0 \end{cases}$$

PROOF Since $\Theta(\tilde{\alpha})$ is half the twist of $\tilde{\alpha}$ at some point in $\mathbb{H}^2$, these bounds follow immediately from previous investigations of section 3.4 into the twist of elements of $\widetilde{PSL_2\mathbb{R}}$. ∎

3.7 Relationship to the trace

Note that $\widetilde{PSL_2\mathbb{R}}$ covers $SL_2\mathbb{R}$ which covers $PSL_2\mathbb{R}$. So an element of $\widetilde{PSL_2\mathbb{R}}$ has a well-defined projection to $SL_2\mathbb{R}$, and hence a well-defined trace. Given that the trace of a hyperbolic isometry is related to its geometry, we might hope that we can extract information relating to Θ from the trace. We cannot say too much however. Since $\Theta(\tilde{\alpha})$ is a twist angle at a particular point in the plane, it will in general bear little relation to the translation distance or rotation angle of α. We can only make the following simple claims.

Lemma 3.7.1 *The traces of* $\mathbf{z}^n$ *and the regions* Hyp_n *and* Par_n *are given as follows:*

$$\mathrm{Tr}\,(\mathbf{z}^n) = (-1)^n \cdot 2$$
$$\mathrm{Tr}\,(\mathrm{Par}_n) = (-1)^n \cdot 2$$
$$\mathrm{Tr}\,(\mathrm{Hyp}_n) = \begin{cases} (2, \infty) & n \text{ even} \\ (-\infty, -2) & n \text{ odd.} \end{cases}$$

PROOF Recall in the previous section we defined the exponential map $\widetilde{\exp} : \mathbb{R} \longrightarrow \widetilde{SO_2\mathbb{R}} \subset \widetilde{PSL_2\mathbb{R}}$ taking θ to a rotation of 2θ about i. Recall further this map projects to $\exp : \mathbb{R} \longrightarrow SO_2\mathbb{R} \subset PSL_2\mathbb{R}$, also taking θ to a rotation of 2θ about i (this time θ is taken modulo 2π). Explicitly

$$\exp(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

Now clearly $\widetilde{\exp}(n\pi) = \mathbf{z}^n$, and hence $\mathbf{z}^n$ projects to the matrix

$$\exp(n\pi) = \begin{bmatrix} (-1)^n & 0 \\ 0 & (-1)^n \end{bmatrix}$$

from which the first claim follows. Now note that the trace of an element of $\widetilde{PSL_2\mathbb{R}}$ is ± 2 if and only if it is a power of $\mathbf{z}$, or parabolic. Further Tr is a continuous function. Now consider the topology of $\widetilde{PSL_2\mathbb{R}}$ (see figure 3.9). The set $\text{Par}_n \cup \{\mathbf{z}^n\}$ (note $\mathbf{z}^0$ is the identity) is connected and the trace of every point in this set is ± 2. Thus Tr is constant on this set, proving the second claim. And as Hyp_n is connected, Tr either takes values in $(-\infty, -2)$ or $(2, \infty)$ on this set. But Hyp_n is bounded by Par_n and $\mathbf{z}^n$ on which $\text{Tr} = (-1)^n \cdot 2$. This proves the final claim. ■

The following result will be very important in subsequent chapters, when we consider representations of commutators into $PSL_2\mathbb{R}$. Similar results are proved in [28], [49] and [62], but we have not seen a directly geometric proof such as this. The style of the proof will be quite similar to results proved in subsequent chapters.

Proposition 3.7.2 *If $g, h \in PSL_2\mathbb{R}$ then (noting $[g,h]$ is well-defined in $\widetilde{PSL_2\mathbb{R}}$)*

$$[g,h] \in \{1\} \cup \left(\bigcup_{n=-1}^{1} \text{Hyp}_n \cup \text{Ell}_n \right) \cup \text{Par}_0 \cup \text{Par}_{-1}^{+} \cup \text{Par}_1^{-}.$$

(Here we take $\text{Ell}_0 = \emptyset$ for convenience.)

PROOF From corollary 3.5.3 we know $|\Theta([g,h])| \leq 3\pi/2$. From corollary 3.6.4, this implies that $[g,h]$ must lie in $\{1, \mathbf{z}, \mathbf{z}^{-1}\}$, Hyp_n for $-1 \leq n \leq 1$, Ell_n for $-2 \leq n \leq 2$, or Par_n with $-1 \leq n \leq 1$. We must exclude the possibilities $\mathbf{z}, \mathbf{z}^{-1}, \text{Par}_1^{+}, \text{Par}_{-1}^{-}, \text{Ell}_2, \text{Ell}_{-2}$. Note that in all these cases, by 3.7.1, $\text{Tr}[g,h] < 2$. So by 3.2.2, g, h are hyperbolic and their axes cross. We will consider the cases $[g,h] \in \{\mathbf{z}\}, \text{Par}_1^{+}, \text{Ell}_2$; the other cases are obviously similar.

Case (i): $[g,h] = \mathbf{z}$. So $[g,h] = 1 \in PSL_2\mathbb{R}$. Thus g, h commute, so the axes of g, h must be identical. We may lift to $\tilde{g}, \tilde{h}$ which flow unit tangent vectors along

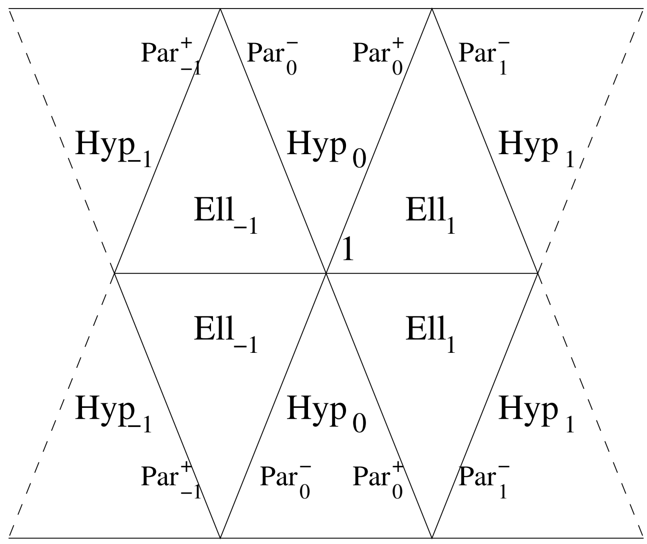

Figure 3.15: Possible commutators in $\widetilde{PSL_2\mathbb{R}}$.

this axis. Then we see $[\tilde{g}, \tilde{h}] = 1$. But by proposition 3.3.1, $[g, h]$ is independent of the choice of lift, so $[g, h] = 1 = \mathbf{z}$, a contradiction.

Case (ii): $[g, h] \in \text{Par}^+_1$**.** We consider the location of the fixed points of $[g, h]$, $[g^{-1}, h]$, $[g, h^{-1}]$, and $[g^{-1}, h^{-1}]$. By lemma 3.2.5, these lie as shown in figure 3.16. (Note the axes of g, h must lie in the arrangement shown; otherwise $[g, h]$ translates clockwise around the circle at infinity, contradicting $[g, h] \in \text{Par}^+_1$.) If q is the fixed point of $[g, h]$ at infinity, note that $h^{-1}q$ is the fixed point of $h^{-1}[g, h]h = [h^{-1}, g]$ at infinity, which is also the fixed point of $[g, h^{-1}]$ at infinity. Similarly $g^{-1}h^{-1}q$ is the fixed point at infinity of $[g^{-1}, h^{-1}]$ and $hg^{-1}h^{-1}q$ is the fixed point at infinity of $[g^{-1}, h]$. So we may consider horocycle neighbourhoods of these fixed points as shown, which are translates of each other under g and h. Take a point p on the horocycle about q.

Again we consider a lift $\tilde{g} \in \widetilde{PSL_2\mathbb{R}}$ of g which translates unit tangent vectors along the axis of g, and along constant distance curves from the axis of g; and similarly for h. Start with a unit tangent vector (p, u) at p. Again consider the action of $\tilde{h}^{-1}, \tilde{g}^{-1}, \tilde{h}, \tilde{g}$ successively upon (p, u). In the situation of the figure shown, the tangent vector is turned clockwise at each stage, along constant distance curves. Thus $\text{Tw}([g, h], p) < 0$. But by lemma 3.4.3, $[g, h] \in \text{Par}^+_1$ implies, $\text{Tw}([g, h], \mathbb{H}^2) = (2\pi, 3\pi)$, a contradiction.

Case (iii): $[g, h] \in \text{Ell}_2$**.** By lemma 3.2.6, the fixed points $p, h^{-1}p, g^{-1}h^{-1}p, hg^{-1}h^{-1}p$ of (respectively) $[g, h], [g, h^{-1}], [g^{-1}, h^{-1}]$ and $[g^{-1}, h]$, lie as shown in figure 3.17. Consider lifts $\tilde{g}, \tilde{h} \in \widetilde{PSL_2\mathbb{R}}$ of g, h as in the previous case. If the axes of g, h are arranged as in figure 3.16 above, then by the same argument as there, $\text{Tw}([g, h], p) <$

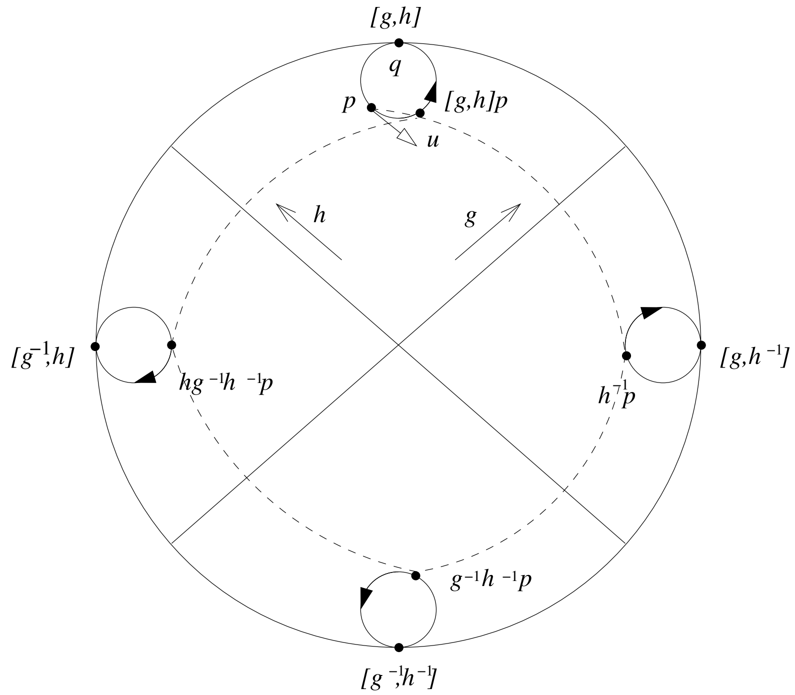

Figure 3.16: Arrangement in case (ii).

0, which is a contradiction as lemma 3.4.4 gives $\mathrm{Tw}([g,h],p) \in (2\pi, 4\pi)$. Thus the situation must be as in figure 3.17.

Consider a unit tangent vector (p,u), and the action of $\tilde{h}^{-1}, \tilde{g}^{-1}, \tilde{h}, \tilde{g}$ upon it. The unit tangent vector is effectively moved by parallel translation along constant distance curves to the axes of g, h. These form a simple quadrilateral bounding an embedded disc in $\mathbb{H}^2$, as shown by the dotted lines. But parallel translation around a simple quadrilateral cannot turn the unit tangent vector by more than a full revolution. So $\mathrm{Tw}([g,h],p) \leq 2\pi$, a contradiction to $\mathrm{Tw}([g,h],p) \in (2\pi, 4\pi)$. ■

Note that since $[g,h] \in \mathrm{Ell}_{\pm 2}$ is excluded, $\Theta([g,h]) = \pm 3\pi/2$ is impossible. Thus $\Theta([g,h]) \in (-3\pi/2, 3\pi/2)$.

Corollary 3.7.3 *If $g, h \in PSL_2\mathbb{R}$ then*

(i) $\mathrm{Tr}[g,h] > 2$ *implies* $[g,h] \in \mathrm{Hyp}_0$*;*

(ii) $\mathrm{Tr}[g,h] = 2$ *implies* $[g,h] \in \{1\} \cup \mathrm{Par}_0$*;*

(iii) $\mathrm{Tr}[g,h] \in (-2,2)$ *implies* $[g,h] \in \mathrm{Ell}_{-1} \cup \mathrm{Ell}_1$*;*

(iv) $\mathrm{Tr}[g,h] = -2$ *implies* $[g,h] \in \mathrm{Par}^+_{-1} \cup \mathrm{Par}^-_1$*;*

(v) $\mathrm{Tr}[g,h] < -2$ *implies* $[g,h] \in \mathrm{Hyp}_{-1} \cup \mathrm{Hyp}_1$.

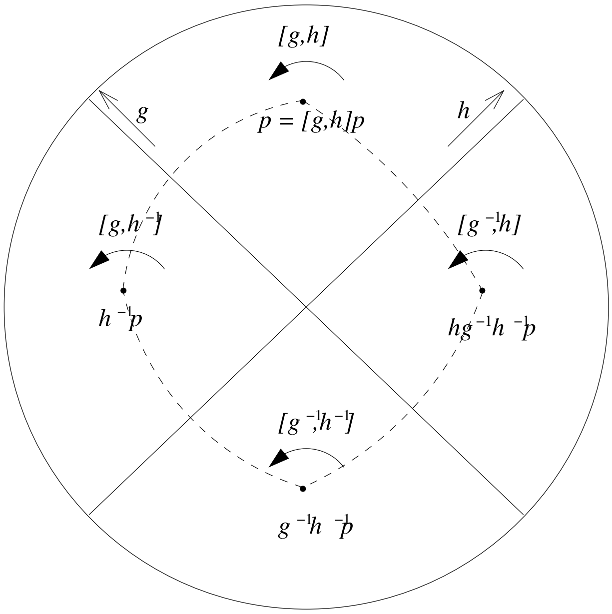

Figure 3.17: Arrangement in case (iii).

PROOF Proposition 3.7.2 tells us that these are the only possible regions of $\widetilde{PSL_2\mathbb{R}}$ to be considered. Then the preceding proposition classifies the regions as shown. ∎

Chapter 4

Euler classes and representation spaces

We now introduce the Euler class, a cohomology class $\mathcal{E}(\rho)$ associated to a representation $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$. We discuss this class in terms of obstruction theory, before proceeding to a discussion of the space of representations, the character variety, and a symplectic structure and measure which can be introduced on it.

4.1 The Euler class on surfaces without boundary

Let S be a closed surface of genus at least 2. As discussed in section 2.1, finding a hyperbolic structure on S is equivalent to finding a transverse section of the bundle $\mathcal{F}(S, \mathbb{H}^2, \rho)$, which is the quotient of $\tilde{S} \times \mathbb{H}^2$ by the action of $\pi_1(S)$, by deck transformations on the first coordinate and via ρ on the second.

We consider a cell complex structure on S, and aim to find a section of $\mathcal{F}(S, \mathbb{H}^2, \rho)$ over the skeleta of S of increasing dimension. A section can trivially be found on the 0-skeleton, sending vertices to arbitrary points in the fibre above them. All such sections are homotopic since $\pi_0(\mathbb{H}^2) = 1$, i.e. $\mathbb{H}^2$ is connected.

We can extend over the 1-skeleton, joining those points (say) by geodesics. This is easier to see if we consider the developing map picture; each edge in $\tilde{S}$ is sent to the geodesic connecting its two endpoints. All such sections are homotopic since $\pi_1(\mathbb{H}^2) = 1$.

This leaves only extension over the 2-skeleton, which is much more difficult. Let σ be a 2-cell of S. We have already constructed the section on $\partial\sigma$, which is homeomorphic to S^1. The map $s|_{\partial\sigma}$ gives a loop in $\mathcal{F}(S, \mathbb{H}^2, \rho)$, and the map $\mathcal{D}$, restricted to a lift $\widetilde{\partial\sigma}$ of $\partial\sigma$, gives a loop in $\mathbb{H}^2$. It is a loop rather than a path, since $\partial\sigma$ is nullhomotopic. If this loop is a simple closed curve in $\mathbb{H}^2$ then we may glue in a disc, which extends the section over σ. But there is no reason in general to expect a simple closed curve. Even if we did obtain simple closed curves around the boundary of each 2-cell, in general they would not be correctly oriented with respect to each other, and folding could occur.

Now consider the fibre bundle $\mathcal{F}(S, PSL_2\mathbb{R}, \rho)$, formed similarly to the previous bundle $\mathcal{F}(S, \mathbb{H}^2, \rho)$. The bundle $\mathcal{F}(S, PSL_2\mathbb{R}, \rho)$ is defined as the quotient of $\tilde{S} \times PSL_2\mathbb{R}$ by the action of $\pi_1(S)$, acting via deck transformations on $\tilde{S}$ and via ρ as isometries (or rather, derivatives of isometries) on $UT\mathbb{H}^2 \cong PSL_2\mathbb{R}$.

We can play the same game on $\mathcal{F}(S, PSL_2\mathbb{R}, \rho)$. We may take a section s_0 over the 0-skeleton arbitrarily, and may extend to s_1 over the 1-skeleton, again by joining points. By choosing s_0 appropriately, we may ensure that s_1 is transverse. However since $\pi_1(PSL_2\mathbb{R}) \cong \mathbb{Z}$, there are many choices for the extension s_1. We may think of extending a developing map by joining points in $\mathbb{H}^2$ by geodesics; but we must also think about unit tangent vectors at each point. Our unit tangent vectors along a geodesic may spin arbitrarily many times.

A global and appropriately transverse section of this fibre bundle would give not only a hyperbolic structure on S, but also a section of the unit tangent bundle on S, i.e. a nowhere vanishing vector field on S. This is quite impossible as the existence of such a vector field requires the Euler characteristic $\chi(S) = 0$; but we are considering hyperbolic surfaces with $\chi(S) < 0$.

Consider a partial section s_1 of $\mathcal{F}(S, PSL_2\mathbb{R}, \rho)$ and partial developing map $\mathcal{D}_1$ defined on the 1-skeleton, and a 2-cell σ in S. On the boundary $\partial\sigma$ (more precisely, a lift $\widetilde{\partial\sigma}$ of $\partial\sigma$) the partial developing map $\mathcal{D}_1$ provides a loop in $UT\mathbb{H}^2$. The free homotopy class of $\mathcal{D}_1(\widetilde{\partial\sigma})$ corresponds precisely to the number of times the unit tangent vector "spins" as it travels around the loop. Returning to the bundle picture, above σ the bundle has local coordinates $\sigma \times PSL_2\mathbb{R}$, so we can consider $s(\sigma)$ as a loop in $PSL_2\mathbb{R}$ by projection to the second coordinate. Over different σ however, the $PSL_2\mathbb{R}$ fibre is shifted by coordinate changes.

Note that $PSL_2\mathbb{R}$ is 1-simple, that is, for $x_1, x_2 \in PSL_2\mathbb{R}$ and any two paths c_1, c_2 from x_1 to x_2, the isomorphisms induced between $\pi_1(PSL_2\mathbb{R}, x_1)$ and $\pi_1(PSL_2\mathbb{R}, x_2)$ are identical. This is clear as $\pi_1(PSL_2\mathbb{R}, x_i) \cong \mathbb{Z}$ is abelian. That is, a change of basepoint affects the fundamental group in a canonical way.

Thus we can consider the coefficient bundle $\mathcal{B}$ associated to $\mathcal{F}(S, PSL_2\mathbb{R}, \rho)$, which is defined formally in [56, 30.2]. This is a $\mathbb{Z}$-bundle over S. We consider $\mathcal{F}(S, PSL_2\mathbb{R}, \rho)$ as a collection of neighbourhoods $V_i \times PSL_2\mathbb{R}$ glued together by coordinate changes $g_{ij} : V_i \times PSL_2\mathbb{R} \longrightarrow V_j \times PSL_2\mathbb{R}$, which is an identification map between open neighbourhoods on the first coordinate and the derivative of an isometry of $\mathbb{H}^2$ on the second coordinate (using $UT\mathbb{H}^2 \cong PSL_2\mathbb{R}$). As $\mathcal{F}(S, PSL_2\mathbb{R}, \rho)$ is flat, g_{ij} is locally constant on the second coordinate and each coordinate change gives a homomorphism $\gamma_{ij} : \pi_1(V_i \times PSL_2\mathbb{R}) \longrightarrow \pi_1(V_j \times PSL_2\mathbb{R})$ where $\gamma_{ij} = g_{ij*}$. These γ_{ij} can equally be considered as maps $\pi_1(PSL_2\mathbb{R}) \longrightarrow \pi_1(PSL_2\mathbb{R})$, and are the transition maps of $\mathcal{B}$.

The coefficient bundle gives us a way of keeping track of elements of $PSL_2\mathbb{R}$ over different parts of S within $\mathcal{F}(S, PSL_2\mathbb{R}, \rho)$. So once we have constructed our 1-section of $\mathcal{F}(S, PSL_2\mathbb{R}, \rho)$, we may assign to each cell the "spin" of our section around it (which is an integer), giving a 2-cochain of S with coefficients in $\mathcal{B}$. It is a cocycle, since it is top-dimensional. Adjustment by a 2-coboundary corresponds to altering the amount of "spin" chosen along each particular edge. So the cohomology class of this 2-cochain does not depend on the choice of 1-section. It can also be seen that this cohomology class does not depend on the cellular decomposition of our surface S chosen. So from ρ we obtain a well-defined *Euler class* $\mathcal{E}(\rho) \in H^2(S; \mathcal{B})$, i.e. $\mathcal{E}(\rho)$ lies in the second cohomology group of S with coefficients in $\mathcal{B}$.

Now suppose that ρ is the holonomy of a hyperbolic cone-manifold structure on S, a closed surface of genus ≥ 2. Let the cone points $p_1, \ldots, p_k$ have orders $s_1, \ldots, s_k$ (recall order s_i is equivalent to cone angle $2\pi(1+s_i)$). We determine the Euler class $\mathcal{E}(\rho)$, and claim it is the Euler characteristic $\chi(S)$ times the fundamental class, with adjustment for cone points — hence the name. One vivid way to see the Euler class is through a vector field on S.

Take a simplicial decomposition of S consisting of geodesic hyperbolic triangles $\sigma_1, \ldots, \sigma_F$. We require that each triangle has a hyperbolic structure without cone points, i.e. all the cone points are vertices of the triangulation. Note that there is a "standard" unit vector field $\mathcal{V}$ on S with precisely one singularity for every vertex, edge and face of S. The orders of the singularities are: 1 on every face; -1 on every edge; and $1 + s_i$ at every vertex, where s_i is the order of the cone point there (possibly zero). See figure 4.1. The sum of the indices of the singularities of $\mathcal{V}$ is then $\chi(S) + \sum s_i$.

Now perturb $\mathcal{V}$ so that the singularities lie off the 1-skeleton. We have a hyperbolic structure on S, hence a section s of $\mathcal{F}(S, \mathbb{H}^2, \rho)$ and a developing map $\mathcal{D}$. Consider the restriction of $\mathcal{D}$ to the 1-skeleton of $\tilde{S}$, and place unit vectors along the edges in accordance with our perturbed $\mathcal{V}$. This gives a 1-section of $\mathcal{F}(S, PSL_2\mathbb{R}, \rho)$. Then the spin of $\mathcal{V}$ around a triangle σ_i is equal to the sum of the indices of singular points of $\mathcal{V}$ inside σ_i, or its negative, depending on whether the orientation induced by $\mathcal{D}$ is the same as the orientation induced by $[S]$. For now assume these orien-

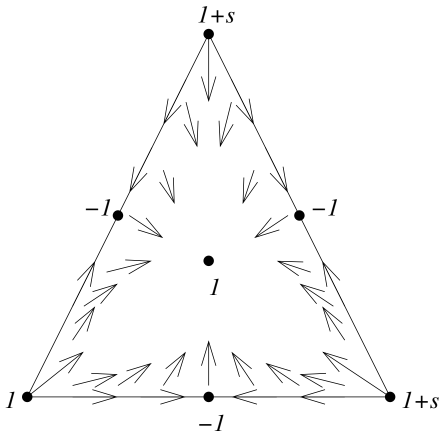

Figure 4.1: The vector field $\mathcal{V}$ on a triangle, with singularities.

tations agree; otherwise all the cohomology classes must be multiplied by -1. (If some triangles folded along an edge, different σ_i would have different orientations. But this does not occur in a hyperbolic structure on S.) Thus $\mathcal{E}(\rho)$ is represented by the 2-cocycle assigning to each σ_i the number of times e_i that $\mathcal{V}$ spins around the boundary of σ_i (represented in the copy of $\pi_1(PSL_2\mathbb{R})$ lying above σ_i in $\mathcal{B}$), as the boundary is traversed according to its orientation. Now $[S] = [\sigma_1] + \cdots + [\sigma_F]$, so $\mathcal{E}(\rho)[S] = \sum e_i$, which is equal (up to sign) to the sum of all indices of singular points in $\mathcal{V}$, which is $\chi(S) + \sum s_i$. We record this conclusion.

Proposition 4.1.1 *Suppose ρ is the holonomy of a hyperbolic cone-manifold structure on a closed surface S with cone points of orders s_i. Then*

$$\mathcal{E}(\rho)[S] = \pm\left(\chi(S) + \sum s_i\right),$$

where $[S]$ is the fundamental class of S. ∎

Note this implies, as a special case, that if ρ is the holonomy of a hyperbolic structure on a closed surface S without cone points, then $\mathcal{E}(\rho)$ takes $[S]$ to $\pm\chi(S)$.

4.2 The relative Euler class

We will now extend the notion of Euler class to surfaces S with boundary, obtaining a *relative Euler class*. It turns out that the relative Euler class can only be defined canonically (i.e. depending only on ρ), when each boundary curve is non-elliptic. More precisely, let $C_1, C_2, \ldots, C_n$ be the boundary curves of S. For a given basepoint $x_0 \in S$ and boundary curve C, we may take a loop C_0 based at x_0 homotopic to C; any two such choices C_0, C_0' are conjugate in $\pi_1(S, x_0)$, and so it makes sense to speak of $\rho(C)$ as elliptic, hyperbolic, parabolic or the identity. We will write $C_i \in \pi_1(S, x_0)$ to denote some chosen loop based at x_0 homotopic to the i'th boundary component. Let $\rho(C_i) = c_i \in PSL_2\mathbb{R}$, and assume that no c_i is elliptic.

We apply a similar procedure. Take a cell decomposition of S and an arbitrary section s_0 of $\mathcal{F}(S, PSL_2\mathbb{R}, \rho)$ on the 0-skeleton. We may extend the section over all 1-cells, as discussed previously, and there are many choices for the homotopy class of the extension over each edge. While there is only one choice of path in $\mathbb{H}^2$ between two points, up to homotopy, the unit tangent vectors may spin an arbitrary number of times along each edge.

In the case of a surface without boundary, it does not matter how we extend over the 1-skeleton, as different choices cancel each other out: each edge belongs to two faces. Similarly here, it does not matter how we extend our section over interior edges. Over boundary edges, however, it does matter.

Here we use the fact that each c_i is not elliptic. Then there is a preferred lift $\tilde{c}_i \in \widetilde{PSL_2}\mathbb{R}$ of c_i, i.e. a preferred "simplest" homotopy class (relative to endpoints) of paths of tangent vectors in $UT\mathbb{H}^2$ along the developing image of the lift of our boundary loop C_i. That is, we have a preferred (small) amount of "spin" we wish to give the unit tangent vectors on this path. We define our 1-section along these edges, or equivalently extend our partial developing map over the universal cover of the 1-skeleton, accordingly. In general, to define a relative Euler class with a boundary curve C_i with $\rho(C_i) = c_i$ elliptic, one also needs to specify a preferred lift $\tilde{c}_i$.

From this section on the 1-skeleton, we may assign a cohomology class $\mathcal{E}(\rho) \in H^2(S; \mathcal{B})$. For each 2-cell σ, $\mathcal{E}(\rho)[\sigma]$ is the number of times our tangent vectors spin as we traverse $\partial\sigma$. Again the cohomology class does not depend on the cell decomposition of S or any of the choices made at any stage. So we obtain a *relative Euler class* $\mathcal{E}(\rho)$ depending only on ρ. If we consider two surfaces joined along a common boundary then we see that the spins along the common boundary cancel out, so that the relative Euler class is additive. This is clearly true, in fact, for any decomposition of a surface into finitely many pieces $S_1, \ldots, S_n$. For the representation on each piece, we must choose a basepoint $p_i \in S_i$, connected to the basepoint p of the overall surface S by a particular path. Thus we may define a representation on $\pi_1(S_i, p_i)$ by restriction and obtain the following result. See also [25], [26].

Lemma 4.2.1 *Suppose a surface S is decomposed along curves C_i, with each $\rho(C_i)$ not elliptic, into surfaces $S_1, S_2, \ldots, S_n$. Then*

$$\mathcal{E}(\rho_1)[S_1] + \cdots + \mathcal{E}(\rho_n)[S_n] = \mathcal{E}(\rho)[S].$$

■

Suppose we have a surface S with boundary, and a hyperbolic cone-manifold structure on S without corner points, but with interior cone points of orders s_i. Then we may consider a vector field $\mathcal{V}$ as in the previous section, and apply the same argument. The same argument applies when each boundary component is totally geodesic; but, as a limiting case, it even goes through when boundary components of S is *cusped*. That is, when a boundary curve C has $\rho(C)$ *parabolic*, the section over C can be taken to be the point at infinity of $\mathbb{H}^2$ fixed by $\rho(C) = c$, and the twist of the simplest lift $\tilde{c} \in \text{Par}_0$ at the fixed point limits to 0, as discussed in section 3.4. We simply take a limiting case of the previous argument. We obtain the following more general result.

Proposition 4.2.2 *Let S be a surface with boundary. Suppose $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ takes every boundary curve to a non-elliptic element. Suppose ρ is the holonomy of a hyperbolic cone-manifold structure on S with no corner points, each boundary component totally geodesic or cusped, and interior cone points of orders s_i. Then the relative Euler class $\mathcal{E}(\rho)$ is well-defined and satisfies*

$$\mathcal{E}(\rho)[S] = \pm \left(\chi(S) + \sum s_i \right).$$

■

4.3 Algebraic description of the Euler class

There is a more directly algebraic way to see the Euler class $\mathcal{E}(\rho)$. Consider the surface S of genus k, with n boundary components, and assume that each c_i is elliptic. This surface is homotopy equivalent to a standard cell structure with one 0-cell, $2k + n$ 1-cells, and one 2-cell, glued as shown in figure 4.2.

We obtain the following standard presentation of the fundamental group $\pi_1(S, x_0)$ (recall composition in π_1 is written *left to right*, while composition of functions like isometries is written *right to left*):

$$\left\langle G_1, H_1, \ldots, G_k, H_k, C_1, \ldots, C_n \mid [G_1, H_1] \cdots [G_k, H_k] C_1 C_2 \cdots C_n = 1 \right\rangle.$$

Consider a lift of the $(4k + n)$-gon fundamental region, lying in $\tilde{S}$. Choose the lift $\tilde{x}_0$ of x_0 as shown in figure 4.2 to be our favourite lift of the baspeoint.

Consider an arbitrary partial section s_0 of $\mathcal{F}(S, PSL_2\mathbb{R}, \rho)$ over the 0-skeleton of S. This consists of a choice of where to place $\tilde{x}_0$ in $\mathbb{H}^2$, a choice of unit tangent vector at that point, and equivariant extension to all lifts of x_0 in $\tilde{S}$. Let $\mathcal{D}(\tilde{x}_0) = (y_0, u_0) \in UT\mathbb{H}^2$.

We now extend to a section s_1 over the 1-skeleton as follows. Draw geodesics between vertices which are joined by an edge in $\tilde{S}$. This gives a system of isometric

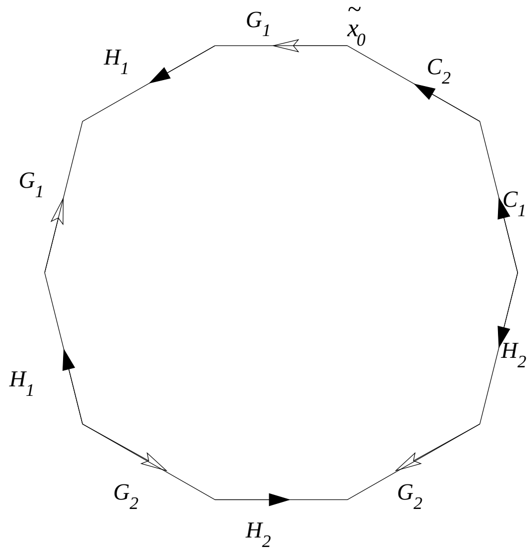

Figure 4.2: Standard cell decomposition of a surface (in this case, $k = 2$, $n = 2$).

$(4k + n)$-gons lying in $\mathbb{H}^2$. A polygon so obtained may however be concave, degenerate or have self-intersections. Letting $\rho(G_i) = g_i$, $\rho(H_i) = h_i$, $\rho(C_i) = c_i$, we arbitrarily choose lifts $\tilde{g}_i, \tilde{h}_i$ of g_i, h_i. Since all the c_i's are non-elliptic we have a preferred lift $\tilde{c}_i$ of each. We now only need choose a path of tangent vectors along each edge.

Let e be an edge and $\tilde{e}$ a lift in the $(4k + n)$-gon in $\tilde{S}$ discussed above. At the endpoints of $\tilde{e}$ a set of tangent vectors is already defined by s_0. These tangent vectors are related by some isometry which is some conjugate of some g_i, h_i or c_i (recall the deck transformation T_{G_i} (for instance), and hence the isometry $g_i = \rho(G_i)$ translates along the edge in $\tilde{S}$ corresponding to G_i which *begins at* $\tilde{x}_0$; translation along a general edge in $\tilde{S}$ corresponding to G_i will be the action of some conjugate of G_i). Taking lifts to $\widetilde{PSL_2\mathbb{R}}$ as described in the previous paragraph gives a unique homotopy class (relative to endpoints) of tangent vectors between these endpoints in $UT\mathbb{H}^2$, which we add to our geodesics. Moving anticlockwise around the polygon in $\tilde{S}$, we now obtain a loop in $UT\mathbb{H}^2$ which is represented by

$$[\tilde{g}_1, \tilde{h}_1] \cdots [\tilde{g}_k, \tilde{h}_k]\ \tilde{c}_1 \cdots \tilde{c}_n$$

(since the holonomy map is a homomorphism — these are composed right to left). By 3.3.1, this element of $\widetilde{PSL_2\mathbb{R}}$ is independent of our choices of lifts $\tilde{g}_i, \tilde{h}_i$. Since $[g_1, h_1] \cdots [g_k, h_k] c_1 \cdots c_n = 1$, this product is $\mathbf{z}^m$, for some $m \in \mathbb{Z}$. This number m is the spin of tangent vectors around the $(4k+n)$-gon. We now have the following.

Proposition 4.3.1 *Let S be an orientable surface with $\chi(S) < 0$. Let $\rho : \pi_1(S, x_0) \longrightarrow PSL_2\mathbb{R}$ be a representation, and let $\pi_1(S)$ have the presentation given above, where* *no c_i is elliptic. The (possibly relative) Euler class $\mathcal{E}(\rho)$ takes the fundamental class $[S]$ to $m \in \mathbb{Z}$ where the unique lift of the relator*

$$[\tilde{g}_1, \tilde{h}_1] \cdots [\tilde{g}_k, \tilde{h}_k]\ \tilde{c}_1 \cdots \tilde{c}_n, \in \widetilde{PSL_2\mathbb{R}}$$

is equal to $\mathbf{z}^m$. ∎

Corollary 4.3.2 *With notation as above, the (possibly relative) Euler class $\mathcal{E}(\rho)$ takes the fundamental class $[S]$ to $m \in \mathbb{Z}$, where $|m| \leq |\chi(S)|$.*

Proof This follows immediately from lemma 3.5.2 and the above. ∎

This result is sometimes known as the *Milnor-Wood inequality*: [49], [62], [28].

Lemma 4.3.3 *For a closed surface S, an abelian representation ρ has $\mathcal{E}(\rho)[S] = 0$. For a general surface S, if the relative Euler class $\mathcal{E}(\rho)$ is well-defined then an abelian representation has $\mathcal{E}(\rho)[S] = 0$.*

Proof We must have each $[g_i, h_i] = 1 \in PSL_2\mathbb{R}$, so $[g_i, h_i] = 1 \in \widetilde{PSL_2\mathbb{R}}$ by proposition 3.7.2: this is enough to dispose of the first assertion by proposition 4.3.1. In a surface with boundary, the c_i now satisfy $c_1 \cdots c_n = 1$, they all commute and are non-elliptic. Hence the c_i consist of hyperbolics with the same axis or parabolics with the same fixed point, along possibly with the identity. Take simplest lifts $\tilde{c}_i$, and consider a unit vector based at some point on the common axis, or on some horocycle. Each $\tilde{c}_i$ acts by translating the unit vector along the common axis or horocycle. Thus they compose to give the trivial path of tangent vectors in $UT\mathbb{H}^2$, and we have $\tilde{c}_1 \cdots \tilde{c}_n = 1 \in \widetilde{PSL_2\mathbb{R}}$. From proposition 4.3.1 then $\mathcal{E}(\rho)[S] = 0$. ∎

Finally, extra information can be extracted from the trace of the commutator in the case of the punctured torus. Take $g = g_1$, $h = g_1$, $c = c_1$. Then $\mathcal{E}(\rho)$ will be well-defined provided that $[g, h]$ is non-elliptic.

Proposition 4.3.4 *If S is a punctured torus, then if the relative Euler class is well-defined:*

(i) $\mathrm{Tr}[g, h] \geq 2$ *is equivalent to* $\mathcal{E}(\rho)[S] = 0$*;*

(ii) $\mathrm{Tr}[g, h] \leq -2$ *is equivalent to* $\mathcal{E}(\rho)[S] = \pm 1$.

Furthermore, in the latter case the positive or negative value is taken, as $[g, h] \in$ $\mathrm{Hyp}_1 \cup \mathrm{Par}_1$ or $\mathrm{Hyp}_{-1} \cup \mathrm{Par}_{-1}$ respectively.

Proof As $\mathcal{E}(\rho)$ is well-defined, $[g,h]$ is not elliptic, hence $\text{Tr}[g,h] \in (-\infty,-2] \cup [2,\infty)$.

Suppose $\text{Tr}[g,h] \geq 2$. Then by corollary 3.7.3 we have $[g,h]$ in $\{1\}$, Hyp_0 or Par_0. We now refer continually to the bounds in corollary 3.6.4. We have $\Theta([g,h]) \in (-\pi/2,\pi/2)$. Taking $\tilde{c} \in \widetilde{PSL_2\mathbb{R}}$ to be the simplest lift of the boundary, we have $[g,h]c = 1$ so $[g,h]$ and c are inverses, hence their angles sum to a multiple of π. But $\tilde{c}$ is a simplest lift and is hyperbolic or parabolic or the identity, so $\Theta(\tilde{c}) \in (-\pi/2,\pi/2)$ and we have $\Theta([g,h])+\Theta(\tilde{c}) \in (-\pi,\pi)$. As this is a multiple of π we have $\Theta([g,h])+\Theta(\tilde{c}) = 0$. Now approximate additivity of Θ 3.5.1 gives $\Theta([g,h]\tilde{c}) \in (-\pi/2,\pi/2)$ but as $[g,h]c = 1$ we must have $[g,h]\tilde{c} = 1 \in \widetilde{PSL_2\mathbb{R}}$. Hence $\mathcal{E}(\rho)[S] = 0$ by proposition 4.3.1 above.

Now suppose $\text{Tr}[g,h] \leq -2$. Again by corollary 3.7.3 we have $[g,h]$ lies in $\text{Hyp}_{\pm 1}$, Par_{-1}^{+} or Par_{1}^{-}. Assume $[g,h] \in \text{Hyp}_1 \cup \text{Par}_1^-$ so $\Theta([g,h]) \in (\pi/2, 3\pi/2)$; the other case is similar. Take a simplest $\tilde{c} \in \widetilde{PSL_2\mathbb{R}}$ such that $[g,h]c = 1 \in PSL_2\mathbb{R}$. As $[g,h]$ is hyperbolic or parabolic, so is $\tilde{c}$ so $\tilde{c} \in \text{Hyp}_0 \cup \text{Par}_0$ and we have $\Theta(\tilde{c}) \in (-\pi/2,\pi/2)$. As above we obtain $\Theta([g,h])+\Theta(\tilde{c}) \in (0,2\pi)$ and is a multiple of π, so equals π. And then approximate additivity 3.5.1 gives $\Theta([g,h]\tilde{c}) \in (\pi/2,3\pi/2)$ and is a multiple of π, so is equal to π, and $\mathcal{E}(\rho)[S] = 1$. The case $[g,h] \in \text{Hyp}_{-1} \cup \text{Par}_{-1}$ similarly implies $\mathcal{E}(\rho)[S] = -1$. ■

4.4 The character variety and its measure

For a general surface S, the *representation variety* $R(S)$ which is the set of all homomorphisms $\rho : \pi_1(S) \longrightarrow SL_2\mathbb{R}$. If we consider a presentation for $\pi_1(S)$, we see that a choice of homomorphism ρ amounts to choosing for each generator of $\pi_1(S)$ a matrix in $SL_2\mathbb{R}$, such that the matrices satisfy the conditions of any relators. The entries of the matrices can be considered as coordinate variables, so that $R(S)$ is the solution set of some polynomial equations. That is, $R(S)$ is a closed algebraic set. The space of representations into $PSL_2\mathbb{R}$ can be obtained by taking an obvious quotient of this space.

For a closed surface S of genus $g \geq 2$, the space $R(S)$ is not connected. In fact if we vary a representation continuously, we see that $\mathcal{E}(\rho)[S]$ changes continuously, but must take an integer value between $\chi(S)$ and $-\chi(S)$; hence it remains constant. In [28] Goldman classified the components of $R(S)$ completely:

Theorem 4.4.1 *For a closed surface S, $R(S)$ has precisely $2|\chi(S)|+1$ components, parameterised by the Euler class.* ■

He also proved in [25], and we shall reprove in chapter 8, that a representation ρ is the holonomy of a complete hyperbolic structure on S if and only if ρ lies in an extremal component, i.e. $\mathcal{E}(\rho)[S] = \pm\chi(S)$.

Again considering a general surface with boundary, the *character* χ of a representation ρ is the function $\chi : \pi_1(S) \longrightarrow \mathbb{R}$ given by $\chi(G) = \mathrm{Tr}(\rho(G))$. By using trace relations such as

$$\mathrm{Tr}\, g^{-1} = \mathrm{Tr}\, g, \quad \mathrm{Tr}\, g^2 = \mathrm{Tr}^2\, g - 2, \quad \mathrm{Tr}\, gh^{-1} = \mathrm{Tr}\, g \,\mathrm{Tr}\, h - \mathrm{Tr}\, gh$$

we can show that the function χ is determined by its values at only finitely many elements $\gamma_1, \ldots, \gamma_m$ of $\pi_1(S)$: see [15]. We can then define a function $t : R(S) \longrightarrow \mathbb{R}^m$ by $t(\rho) = (\mathrm{Tr}(\rho(\gamma_1)), \ldots, \mathrm{Tr}(\rho(\gamma_m)))$ and define the *character variety* to be $X(S) = t(R(S))$. It can be shown that $X(S)$ is a closed algebraic set: again see [15].

There is an action of $\pi_1(S)$ on $R(S)$ by conjugation, and we can consider the quotient space $R(S)/\pi_1(S)$. We can think of this quotient space as the moduli space of isomorphism classes of flat principal $SL_2\mathbb{R}$-bundles over S. In general it has singularities. The character variety can be considered as an "algebraic" version of this quotient. Away from singularities, the character variety and this quotient can be identified.

There is a symplectic structure on this quotient $R(S)/\pi_1(S)$ — that is, a closed non-degenerate 2-form — although the structure is singular along the singularities of $R(S)$ ([29]). To see how this structure arises, we need to consider the tangent spaces to $R(S)$ and $R(S)/\pi_1(S)$. Since these are algebraic sets, we consider the Zariski tangent space, which has an interpretation in terms of group cohomology. We briefly describe it now: see [33] or [26] for more details.

Consider a smooth path ρ_t of representations in $R(S)$. We can then approximate ρ_t to first order:

$$\rho_t(G) = \exp\left(tu(G) + O(t^2)\right) \rho_0(G),$$

where exp is the exponential map for the Lie group $SL_2\mathbb{R}$ and u is some function from $\pi_1(S)$ into the Lie algebra $\mathfrak{sl}_2\mathbb{R}$. To first order ρ_t is determined by u. The fact that each ρ_t is a homomorphism implies that for each t and all $G, H \in \pi_1(S)$, we have $\rho_t(GH) = \rho_t(G)\rho_t(H)$. This is equivalent to

$$u(GH) = u(G) + \mathrm{Ad}(\rho_0(G))\ (u(H)).$$

Here $\mathrm{Ad} : SL_2\mathbb{R} \longrightarrow \mathrm{Aut}(\mathfrak{sl}_2\mathbb{R})$ is the adjoint representation. Now the adjoint, combined with the representation ρ_0, defines an $\mathbb{R}\pi_1(S)$-module structure on $\mathfrak{sl}_2\mathbb{R}$, given by $G.v = \mathrm{Ad}(\rho_0(G))(v)$, for $G \in \pi_1(S)$ and $v \in \mathfrak{sl}_2\mathbb{R}$. We denote this $\mathbb{R}\pi_1(S)$-module by $\mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0}$. The condition that u must satisfy is then just the condition that $u : \pi_1(S) \longrightarrow \mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0}$ is a *1-cocycle* in the group cohomology of $\pi_1(S)$ with coefficients in $\mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0}$, i.e. $u \in Z^1(\pi_1(S); \mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0})$ (see e.g. [9] or [32] for details). Thus the Zariski tangent space to $R(S)$ at ρ_0 can be identified with the $\mathbb{R}$-vector space structure on this cocycle module.

Now we consider the tangent space of $R(S)/\pi_1(S)$, the quotient space by conjugation. Thus there is a map on tangent spaces $T_{\rho_0}R(S) \longrightarrow T_{[\rho_0]}\left(R(S)/\pi_1(S)\right)$. A path $\rho_t \in R(S)$ of representations corresponds to a tangent vector in the kernel of this quotient if and only if, to first order, each ρ_t is conjugate to ρ_0, i.e. $\rho_t(x) = g_t^{-1}\rho_0(x)g_t$ for some path $g_t \in SL_2\mathbb{R}$. So let $g_t = \exp(tu_0 + O(t^2))$ and again let $\rho_t = \exp\left(tu(x) + O(t^2)\right)\rho_0$. The condition that u be in the kernel is precisely the coboundary condition

$$u(x) = \mathrm{Ad}(\rho_0(x))(u_0) - u_0 = \delta u_0 \in B^1\left(\pi_1(S); \mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0}\right).$$

So the tangent space to $R(S)/\pi_1(S)$ at $[\rho_0]$ is the (vector-space structure on the) cohomology module $H^1(\pi_1(S); \mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0})$.

Restricting our attention to a closed surface of genus g, we can make calculations of the dimensions of these spaces: see [26]. The dimension of the tangent space to $R(S)$ at ρ_0 is $\dim Z^1(\pi_1(S); \mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0}) = 6g - 3 + \dim C(\rho_0)$, where $C(\rho_0)$ is the centraliser of $\rho_0(\pi_1(S))$ in $SL_2\mathbb{R}$. We see $C(\rho_0)$ is trivial for non-abelian ρ_0, 1-dimensional for non-trivial abelian ρ_0, and all of $SL_2\mathbb{R}$ (hence 3-dimensional) for $\rho_0 = 1$. Thus for all non-abelian ρ_0, the tangent space $T_{\rho_0}R(S)$ has dimension $6g - 3$. And $\dim B^1(\pi_1(S); \mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0}) = 3 - \dim C(\rho_0)$. Thus the dimension of the tangent space to $R(S)/\pi_1(S)$ at $[\rho_0]$ is $\dim H^1(\pi_1(S); \mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0}) = 6g - 6 + 2C(\rho_0)$. Letting $R(S)^-$ denote the non-abelian representations, we may take the quotient $R(S)^-/\pi_1(S)$, which is $(6g-6)$-dimensional. In general however this space is not Hausdorff: [26]. The characters of abelian representations are precisely the singularities of $R(S)/\pi_1(S)$.

Returning to a general surface with boundary, we consider the cup product in group cohomology on $\pi_1(S)$ with coefficients in $\mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0}$. This gives a dual pairing

$$H^1\left(\pi_1(S); \mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0}\right) \times H^1\left(\pi_1(S); \mathfrak{sl}_2\mathbb{R}^*_{\mathrm{Ad}\,\rho_0}\right) \longrightarrow H^2(\pi_1(S); \mathbb{R}) = \mathbb{R}.$$

There is a nondegenerate symmetric bilinear form (the Killing form) on $\mathfrak{sl}_2\mathbb{R}$, which is invariant under the adjoint representation, giving an isomorphism $\mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0} \cong \mathfrak{sl}_2\mathbb{R}^*_{\mathrm{Ad}\,\rho_0}$. Using this form with the cup product we can define a dual pairing on $R(S)/\pi_1(S)$

$$\omega_{\rho_0} : \; H^1\left(\pi; \mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0}\right) \times H^1\left(\pi; \mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_0}\right) \longrightarrow \mathbb{R}.$$

Since these cohomology groups are identified with the tangent space to $R(S)/\pi_1(S)$, ω_{ρ_0} is a 2-form on $T_{\rho_0}(R(S)/\pi_1(S))$ (i.e. on the tangent space at ρ_0). This clearly varies continuously with $[\rho_0]$, so we obtain a 2-form ω on $R(S)/\pi_1(S)$, which is singular at the equivalence classes of abelian representations. It can be shown (see [26]) that ω is closed and nondegenerate.

If S is a closed surface, then $R(S)/\pi_1(S)$ is everywhere even-dimensional (even though the dimension varies) and we obtain a symplectic structure on $R(S)/\pi_1(S)$. Hence we obtain a symplectic structure on $X(S)$, away from the characters of abelian representations. By taking an appropriate exterior power of ω, we obtain an area form on $R(S)^-/\pi_1(S)$, and a singular area form on $R(S)/\pi_1(S)$. This gives a measure on $R(S)/\pi_1(S)$. It can be shown that the singular set has measure zero and the measure of $R(S)/\pi_1(S)$ is finite: see [29], [37]. So we may obtain a measure μ_S on $X(S)$. Considering $X(S)$ as a subset of some $\mathbb{R}^{2N}$, away from singular points ω^N is some multiple of the standard Euclidean area form, hence μ_S is absolutely continuous with respect to Lebesgue measure.

The character variety can be defined in a similar way for any manifold. In particular it is clear that for a circle S^1 we obtain $X(S^1) \cong \mathbb{R}$. At points in $\mathbb{R}$ other than ± 2 the character defines the conjugacy class of a representation uniquely.

For a surface S with boundary, we can consider a *relative character variety*, following [29]. The boundary ∂S is a collection of circles $C_1, \ldots, C_n$, and so we obtain $X(\partial S) = X(S^1)^n = \mathbb{R}^n$. There is then a restriction map

$$\partial^{\#} : X(S) \longrightarrow X(\partial S) = \mathbb{R}^n.$$

If we specify for each C_i a conjugacy class $\mathcal{C}_i$, then we may define the *relative character variety* to be

$$X_{\mathcal{C}}(S) = \{[\rho] \in R(S)/\pi_1(S) \mid \rho(C_i) \in \mathcal{C}_i\}.$$

Note that if $\mathcal{C}_i$ is hyperbolic or elliptic, then it is described completely by its trace, and we can write $X_t(S)$.

4.5 The action on the character variety

We now consider the effect of changing a representation $\rho : \pi_1(S) \longrightarrow SL_2\mathbb{R}$ by pre-composition with an automorphism of $\pi_1(S)$: that is, take $\phi \in \operatorname{Aut} \pi_1(S)$ and replace ρ with $\rho' = \rho \circ \phi$. In a sense, applying an automorphism in this way should change nothing in terms of the underlying geometry: the image of ρ is identical, and we have merely changed our presentation of ρ. But the representation ρ' may certainly have a different character $\chi' = \operatorname{Tr} \circ \rho'$ to the character $\chi = \operatorname{Tr} \circ \rho$ of ρ.

There is therefore an action of $\operatorname{Aut}\pi_1(S)$ on the character variety. Since the action by conjugations $\operatorname{Inn}\pi_1(S)$ is trivial (traces are invariant under conjugation), we can also consider the action of the quotient $\operatorname{Out}\pi_1(S) = \operatorname{Aut}\pi_1(S)/\operatorname{Inn}\pi_1(S)$. Points in $X(S)$ which are related under this action ought to be considered as equivalent in terms of the underlying geometry. In chapter 6, we will find precisely which elements of $X(S)$ are equivalent under this action, in the case of a punctured torus.

When S is a closed surface, $\operatorname{Out}\pi_1(S)$ has a geometric interpretation. Every homeomorphism of S which preserves a basepoint determines an automorphism of $\pi_1(S)$. A general homeomorphism of S determines an automorphism of $\pi_1(S)$, up to conjugacy: that is, an outer automorphism. On the other hand, for closed surfaces, the Dehn–Nielsen theorem (see e.g. [57], [52]) states that every automorphism of the fundamental group is induced by a homeomorphism of the surface. In chapter 6 we will see that this remains true for the once-punctured torus; but it is not true for any other surface with boundary.

Now homeomorphisms which are isotopic can be taken as equivalent, and determine conjugate automorphisms of $\pi_1(S)$. Conversely, automorphisms which are conjugate can be taken as equivalent, and determine isotopic homeomorphisms. Thus for S any closed surface or the punctured torus, there is an isomorphism

$$\operatorname{MCG}(S) = \frac{\operatorname{Homeo}(S)}{\text{Isotopy}} \cong \frac{\operatorname{Aut}\pi_1(S)}{\operatorname{Inn}\pi_1(S)} = \operatorname{Out}\pi_1(S).$$

Here $\operatorname{MCG}(S)$ is the mapping class group of S, i.e. the group of homeomorphisms up to isotopy.

The 2-form ω is invariant under the action of $\operatorname{Out}\pi_1(S) \cong \operatorname{MCG}(S)$ on $X(S)$. Recall that the 2-form ω arose as a bilinear pairing on each tangent space $T_{\rho_0}\left(R(S)/\pi_1(S)\right)$, coming from the cup product on the cohomology of $\pi_1(S)$ with coefficients in $\mathfrak{sl}_2\mathbb{R}_{\operatorname{Ad}\rho}$. Since S is a $K(\pi_1(S),1)$ space, we have that

$$H^1\left(\pi_1(S), \mathfrak{sl}_2\mathbb{R}_{\operatorname{Ad}\rho_0}\right) \cong H^1(S;\mathcal{B})$$

where $\mathcal{B}$ is the bundle of coefficients over S associated with the $\pi_1(S)$-module $\mathfrak{sl}_2\mathbb{R}_{\operatorname{Ad}\rho_0}$: see [9, p. 59], [17]. Take $\phi \in \operatorname{Aut}\pi_1(S)$ and let $\rho_0' = \rho_0 \circ \phi$ and $\mathcal{B}'$ the bundle associated with $\mathfrak{sl}_2\mathbb{R}_{\operatorname{Ad}\rho_0'}$. There is a corresponding basepoint-preserving homeomorphism $f \in \operatorname{Homeo}(S)$, and we see that $\mathcal{B}'$ is isomorphic to $\mathcal{B}$ via f. Thus $H^1(S;\mathcal{B}') \cong H^1(S;\mathcal{B})$. Under this isomorphism the cup product in the homology of the surface S is preserved, hence the 2-form ω is preserved. So ω is invariant under the action of $\operatorname{Out}\pi_1(S)$, and the action of $\operatorname{Out}\pi_1(S)$ is measure-preserving with respect to μ_S.

Chapter 5

The Geometry of Punctured Tori

In this chapter let S denote a punctured torus with a hyperbolic cone-manifold structure. From Gauss-Bonnet (proposition 2.2.2) we have

$$\sum s_i < -\chi(S) = 1,$$

where for interior cone points $\theta_i = 2\pi(1 + s_i)$ and for corner points $\theta_i = 2\pi(\frac{1}{2} + s_i)$. Since we are only interested in cone points with angles which are multiples of 2π, such s_i are positive integers; hence we cannot have any interior cone points. We allow S to have at most one corner point p_0, with corner angle θ. Then $s_i < 1$ implies $\theta \in (0, 3\pi)$.

The majority of this chapter (particularly section 5.2) is intended to give the reader a feel for the nature of such geometric structures. Indeed the reader will note the prevalence of pictures to describe the situation. The bulk of the proof of the existence of these structures is delayed to the following chapters.

5.1 Pentagon decomposition

We will demonstrate a standard, almost canonical, decomposition of S. This will require some general results about short geodesics in hyperbolic cone-manifolds. Since a regular point has a neighbourhood isometric to a ball in $\mathbb{H}^2$, the behaviour of geodesics at regular points is well-understood. We need to be careful, however, with cone and corner points.

We have seen in the proof of the Hopf-Rinow theorem that given any $p, q \in S$, there is a geodesic joining them, which is a shortest curve between p and q. We now need a little more detail.

Lemma 5.1.1 *Let S be a hyperbolic cone-manifold, and let $p, q \in S$. Then a shortest geodesic C between p and q has the following properties:*

(i) C is simple, i.e. non-self-intersecting, and

(ii) if C intersects the boundary ∂S then $C \cap \partial S$ is a disjoint union of closed segments whose endpoints are corner points with corner angles greater than π, or p or q.

Proof If C intersects itself then we may remove the loop between two points of self-intersection and obtain a shorter geodesic between p and q, contradicting the minimality of C.

If the interior of C intersects ∂S, then since ∂S is closed and C is simple $C \cap \partial S$ is a disjoint union of closed segments. Now p and q may lie on ∂S, but consider a point $w \neq p, q$ where C enters or leaves the boundary. The point w cannot be a regular boundary point, since the angle formed by C at w would then be less than π. The same argument shows w cannot be a corner point with corner angle less than π. Thus w must be a corner point with corner angle greater than π. ∎

We also need the following lemma. Recall that a curve C is *boundary-parallel* to a boundary component A if C can be homotoped to lie entirely on A. In particular a null-homotopic curve is boundary-parallel to A.

Lemma 5.1.2 *Suppose S has no interior cone points and a boundary component A with exactly one corner point q. Then there is a shortest closed curve C based at q which is not boundary-parallel to A. The curve C is a geodesic arc and is simple.*

Proof Since a sufficiently small neighbourhood of q is contractible, the quantity

$$d = \inf \{l(\gamma) : \gamma \text{ is based at } q \text{ and not boundary-parallel to } A\}$$

is positive. Thus we find curves γ_n based at q not homotopic to A such that $l(\gamma_n) < d + 1/n$. As in the proof of Hopf-Rinow, we apply the Arzelà-Ascoli theorem to find a subsequence of γ_n converging uniformly to a curve C with $l(C) = d$.

We claim C is homotopic to γ_n for n sufficiently large, following [8, prop I.3.16], and hence C is not boundary-parallel to A. Take $\epsilon > 0$ sufficiently small so that every closed curve of length less than ϵ is nullhomotopic. By uniform convergence then there exists $N \in \mathbb{N}$ such that, for all $t \in S^1$, $n > N$ implies $d(\gamma_n(t), C(t)) < \epsilon/4$. For a given $n > N$, choose sufficiently many points $\theta_0, \theta_1, \ldots, \theta_m$ on S^1 (where $\theta_0 = \theta_m$) such that each $d(\gamma_n(\theta_i), \gamma_n(\theta_{i+1})) < \epsilon/4$ and $d(C(\theta_i), C(\theta_{i+1})) < \epsilon/4$.

Let p_i be a curve of length less than $\epsilon/4$ joining $\gamma_n(\theta_i)$ and $C(\theta_i)$. Consider the loop obtained by traversing p_i, then $C|_{[\theta_i, \theta_{i+1}]}$, then p_{i+1} backwards, then $\gamma_n|_{[\theta_i, \theta_{i+1}]}$ backwards. This has length less than ϵ, so is nullhomotopic. Since each of these loops is nullhomotopic (see figure 5.1), C is homotopic to γ_n as required.

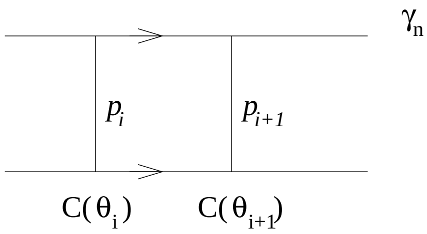

Figure 5.1: C is homotopic to γ_n.

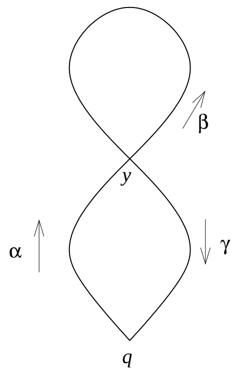

Figure 5.2: If C is not simple.

So C is based at q, C is not boundary-parallel to A, and $l(C) = d$, the minimal length of all such curves.

Since q is the only singular point of S, the curve C can only consist of geodesic arcs from q to q which avoid the corner point q in their interior. These segments either lie entirely in the interior of S, or lie entirely on the boundary component A. If there is a segment lying along the boundary component A, then removing this segment shortens C and C remains non-boundary-parallel to A (if C were now boundary-parallel, then C was originally boundary parallel). Similarly, if there is a boundary-parallel segment lying in the interior of S then this can be removed, shortening C and retaining the non-boundary-parallel character of C. Both these contradict the minimality of C.

Thus C consists of non-boundary-parallel geodesic arcs based at q whose interiors lie in the interior of S. If there is more than one such arc then we may remove it, contradicting the minimality of C. So C is a single geodesic arc and intersects the boundary only at q, at its endpoints.

Suppose C is not simple. Then C intersects itself at some point y in the interior of S. We denote the three segments $q \to y \to y \to q$ by α, β, γ respectively, so $l(C) = l(\alpha) + l(\beta) + l(\gamma)$. See figure 5.2. The intersection at y must be transverse: y must be a regular point, and a geodesic through a regular point has a unique continuation. So if the intersection were not transverse, the geodesic segments would coincide.

If the curve $\alpha.\gamma$ (writing composition from left to right, so $c = \alpha.\beta.\gamma$) is not boundary parallel to A, then we have a contradiction to the minimality of C. So $\alpha.\gamma$ is boundary parallel. Thus the curve $\alpha.\beta.\alpha^{-1}$ is not boundary parallel, and the curve β (considered as a free loop subject to free homotopy) is not boundary parallel to A. (If it were, then the composition $(\alpha.\beta.\alpha^{-1}).(\alpha.\gamma) = \alpha.\beta.\gamma = C$ would be boundary parallel also.) Similarly the curve $\gamma.\beta.\gamma^{-1}$ is not boundary parallel.

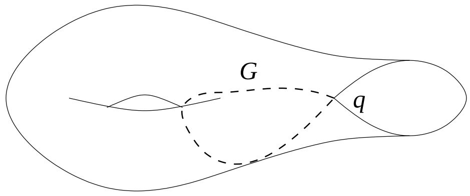

Figure 5.3: Cutting along G

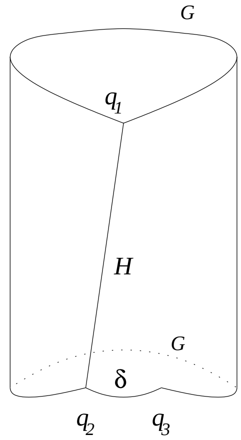

Figure 5.4: Cutting along H

The minimality of C then implies that

$$l(\alpha) + l(\beta) + l(\gamma) \leq 2l(\alpha) + l(\beta), \qquad l(\alpha) + l(\beta) + l(\gamma) \leq l(\beta) + 2l(\gamma),$$

hence $l(\gamma) \leq l(\alpha) \leq l(\gamma)$, so $l(\alpha) = l(\gamma)$. Then the curve $\alpha.\beta.\alpha^{-1}$ is a curve with the same length as C, which is not boundary parallel but is not a geodesic, since the angle at y between the end of β and the start of α^{-1} is not π. Thus C can be shortened, contradicting the minimality of C. ∎

Now we obtain our decomposition.

Proposition 5.1.3 *Let S be a punctured torus with a hyperbolic cone-manifold structure with no interior cone points and at most one corner point q with corner angle θ (let $\theta = \pi$ if q is a regular point). There exist two closed geodesic arcs G, H on S based at q such that cutting along G and H produces a topological disc which is isometric to a geodesic pentagon in $\mathbb{H}^2$ bounding an immersed open disc.*

Proof Let G denote the shortest closed curve through p_0 which is not boundary-parallel. This is guaranteed by lemma 5.1.2, which shows that G is a geodesic arc and is simple.

We cut along G, forming a cylinder with a hyperbolic cone-manifold structure. The two boundary components ∂_1, ∂_2 are piecewise geodesic. There is one corner point q_1 on ∂_1, and two corner points q_2, q_3 on ∂_2. (These were all identified to q in the original S.)

Now consider the shortest curve γ from q_1 to q_2. This curve is piecewise geodesic, with possible corners at q_1, q_2, q_3. It cannot pass through q_1 other than at the start, because of its minimality. Nor can it pass through q_2 other than at the end. If

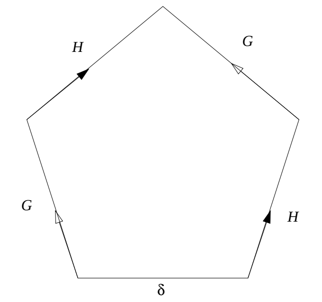

Figure 5.5: The pentagon $\mathcal{P}$ bounds an immersed open disc

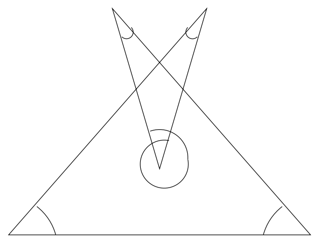

Figure 5.6: $\mathcal{P}$ need not be a simple pentagon. In this case $\mathcal{P}$ still bounds an immersed open disc.

it consists of one geodesic segment from q_1 to q_2, then we let this curve be H. Otherwise γ passes through q_3 on the way to q_2; in this case we take H to be the initial segment from q_1 to q_3.

Thus we obtain a geodesic H on S which intersects G only at q. Cutting along G, H reduces S to a topological disc. Since G, H are geodesic arcs, the developing map of a lift of this topological disc shows that the obtained surface is isometric with an immersed open disc in $\mathbb{H}^2$ bounded by a geodesic pentagon. ∎

Let us now consider the pentagon $\mathcal{P}$ obtained by this procedure. It is certainly non-degenerate, i.e. all of its sides have nonzero length. Two pairs of sides are identified, which correspond respectively to the curves G and H along which we have cut. The sum of the interior angles of the pentagon is equal to the corner angle θ at q. Furthermore, $\{G, H\}$ forms a free basis for $\pi_1(S, q)$ and we have a presentation $\pi_1(S, q) = \langle G, H \rangle$. In the fundamental group, the boundary of S is represented by the commutator $[G, H]$. Note that $\mathcal{P}$ need not be a simple pentagon, if some of the interior angles of the pentagon are large. See figure 5.6.

Now the universal cover $\tilde{S}$ can be considered as a tessellation by copies of this pentagon according to the edge pairings. The developing map of the cone-manifold structure on S is a (generally overlapping) tessellation by isometric copies of the pentagonal fundamental domain $\mathcal{P}$ in $\mathbb{H}^2$.

Let $\rho : \pi_1(S, q) \longrightarrow PSL_2\mathbb{R}$ be the holonomy map. Choose a basepoint $\tilde{q}$ lifting q

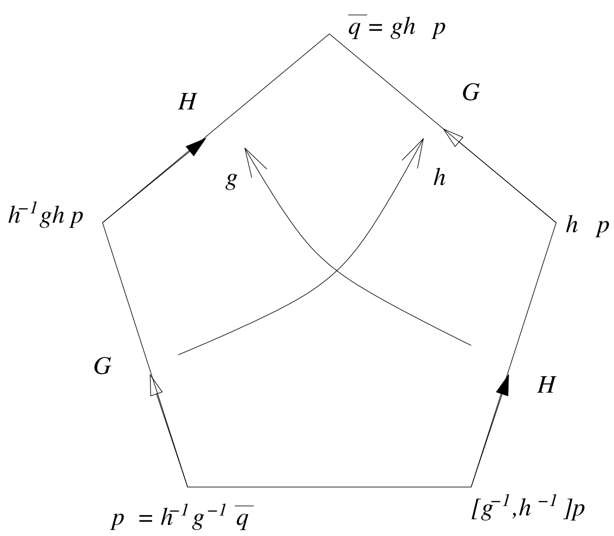

Figure 5.7: Edge identifications in $\mathcal{P}$

in $\tilde{S}$, and its developing image $\bar{q} \in \mathbb{H}^2$. Let $\rho(G) = g$, $\rho(H) = h$. Then with $\bar{q}$ and $\mathcal{P}$ as shown in figure 5.7, g and h identify pairs of sides in $\mathcal{P}$: the isometry g identifies the pair of sides corresponding to the curve H, and the isometry h identifies the pair of sides corresponding to the curve G, as shown.

Labelling one of the vertices $p = h^{-1}g^{-1}\bar{q}$ as shown, we can describe the other vertices of $\mathcal{P}$ as the images of q under various combinations of g and h. The pentagon $\mathcal{P}$ can be described as the sequence of segments (composition always written right to left)

$$p \longrightarrow h^{-1}ghp \longrightarrow ghp \longrightarrow hp \longrightarrow g^{-1}h^{-1}ghp \longrightarrow p.$$

Thus a hyperbolic cone-manifold structure on S with no interior cone points and at most one corner point gives rise to a basis G, H of the fundamental group and a holonomy representation ρ such that the pentagon described above is the boundary of an immersed open disc.

Conversely, suppose we are given a representation $\rho : \pi_1(S, q) \longrightarrow PSL_2\mathbb{R}$ and we can find a basis G, H of $\pi_1(S, q)$ and a point $p \in \mathbb{H}^2$ such that the pentagon described above is non-degenerate and bounds an immersed open disc. Then it is clear that this pentagon is a fundamental domain of a developing map for a hyperbolic cone-manifold structure on S with no interior cone points and at most one corner point, with holonomy ρ. The rest of the developing map is obtained by applying the isometries g, h and their inverses repeatedly to $\mathcal{P}$. We record this fact.

Definition 5.1.4 *Let $g, h \in PSL_2\mathbb{R}$ and $p \in \mathbb{H}^2$. Then the geodesic pentagon in $\mathbb{H}^2$ obtained by joining the segments*

$$p \longrightarrow h^{-1}ghp \longrightarrow ghp \longrightarrow hp \longrightarrow g^{-1}h^{-1}ghp \longrightarrow p$$

is called the pentagon generated by g, h at p *and is denoted* $\mathcal{P}(g, h, p)$.

Lemma 5.1.5 *Let $\rho : \pi_1(S, q) \longrightarrow PSL_2\mathbb{R}$ be a representation. The representation ρ is the holonomy of a hyperbolic cone manifold structure on S with no interior cone points and at most one corner point if and only if there exists a free basis G, H of $\pi_1(S, q)$ and a point $p \in \mathbb{H}^2$ such that $\mathcal{P}(g, h, p)$ is a non-degenerate pentagon bounding an immersed open disc in $\mathbb{H}^2$.* ■

Despite its simplicity, lemma 5.1.5 will be crucial in constructing geometric structures with prescribed holonomy.

We will always consider a surface with a hyperbolic cone-manifold structure to have orientation induced from the orientation on the hyperbolic plane. The boundary of S may be oriented according to the direction of $[G, H]$ (composition in $\pi_1(S)$ is left to right, and composition of isometries right to left). In the situation of figure 5.7 for instance, if we traverse the boundary in the direction of $[G, H]$ then the surface lies to our left; similarly, if we traverse the segment $p \to [g^{-1}, h^{-1}]p$ (composition in $PSL_2\mathbb{R}$ is right to left), the pentagon lies to our left.

5.2 Non-uniqueness of geometric structures

There is a great deal of non-rigidity in these hyperbolic cone-manifold structures. For a given representation ρ, there may be many non-isometric structures on S, and even more non-isometric pentagons $\mathcal{P}(g, h, p)$. Isometric pentagons of course give isometric punctured tori, and isometric punctured tori have equivalent (conjugate) holonomy representations. But non-isometric punctured tori can have the same (or conjugate) holonomy representation; and non-isometric pentagons may give isometric punctured tori. We now explain some of these phenomena and give an overview of the different behaviours possible.

The pentagon $\mathcal{P}(g, h, p)$ is the object containing the most information: not only does it encode a hyperbolic cone-manifold structure on the punctured torus, it also encodes a choice of basis curves G, H, and a particular location in $\mathbb{H}^2$. The hyperbolic cone-manifold structure on the torus S encodes less information: it does not encode any choice of basis curves, but it does include particular locations in $\mathbb{H}^2$ via its developing map. Here we take the view that a hyperbolic cone-manifold structure is a particular developing map, rather than an equivalence class of developing maps determined up to isometry. The representation ρ encodes the least information: it determines no basepoint from which to begin a developing map or a pentagon; and no choice of basis.

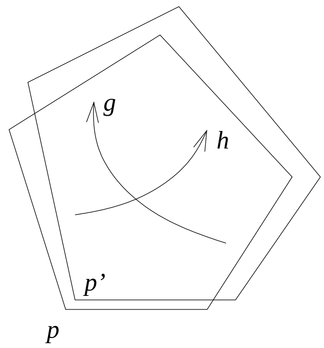

Figure 5.8: $\mathcal{P}(g,h,p)$ and $\mathcal{P}(g,h,p')$ need not be isometric

The diagram below illustrates the situation schematically: solid arrows denote a complete determination of one object by another; broken arrows denote that some choice is involved. Some non-rigidity is thus immediately clear.

$$\left\{ \begin{array}{c} \text{Pentagon} \\ \mathcal{P}(g,h,p) \end{array} \right\} \quad \begin{array}{c} \rightarrow \\ \leftarrow\text{-}\text{-} \\ \text{choose} \\ \text{basis} \end{array} \quad \left\{ \begin{array}{c} \text{Hyp. cone} \\ \text{manifold} \\ \text{structure on } S \\ \text{(developing map)} \end{array} \right\} \quad \begin{array}{c} \rightarrow \\ \leftarrow\text{-}\text{-} \\ \text{choose} \\ \text{basepoint} \end{array} \quad \left\{ \begin{array}{c} \text{Rep.} \\ \rho \end{array} \right\}$$

Let us first consider the relationship between the representation ρ and the hyperbolic cone-manifold structure on the punctured torus S. We investigate, for given ρ, the effect of a choice of different choices of basepoint p from which to generate a developing map. Let g, h be some choice of basis, which we fix for now. We consider a pentagon $\mathcal{P}(g,h,p)$ bounding an immersed open disc, and the effect of perturbing p. With a small perturbation of p to p', the pentagon $\mathcal{P}(g,h,p')$ will still bound an immersed disc. But the two pentagons are in general not isometric; nor are the punctured tori. One easy way to see this is to note that the sides of the pentagons corresponding to ∂S will in general not have the same length. See figure 5.8.

However, still for fixed g, h, different choices of basepoint may give non-isometric pentagons $\mathcal{P}(g,h,p)$, $\mathcal{P}(g,h,p')$ which correspond to isometric cone-manifold structures on S. As the simplest example, suppose ρ is a discrete representation which is the holonomy of a complete hyperbolic structure on S with totally geodesic boundary, which we will denote S_0. We will see that these are precisely the representations with $\text{Tr}[g,h] =< -2$ (note $[g,h]$ and $[g^{-1},h^{-1}]$ are conjugate so $\text{Tr}[g,h] = \text{Tr}[g^{-1},h^{-1}]$). Then the complete hyperbolic surface S_0 is the quotient of a convex subset of $\mathbb{H}^2$ by the image of ρ, which is a discrete subgroup of $PSL_2\mathbb{R}$. This convex subset is the *convex core* of ρ and depends only on ρ, not on g, h or p. Thus the underlying manifold is independent of the choice of g, h. Infinitely many choices of p — namely any p on the axis of $[g^{-1},h^{-1}]$ — give rise to pentagons $\mathcal{P}(g,h,p)$ which

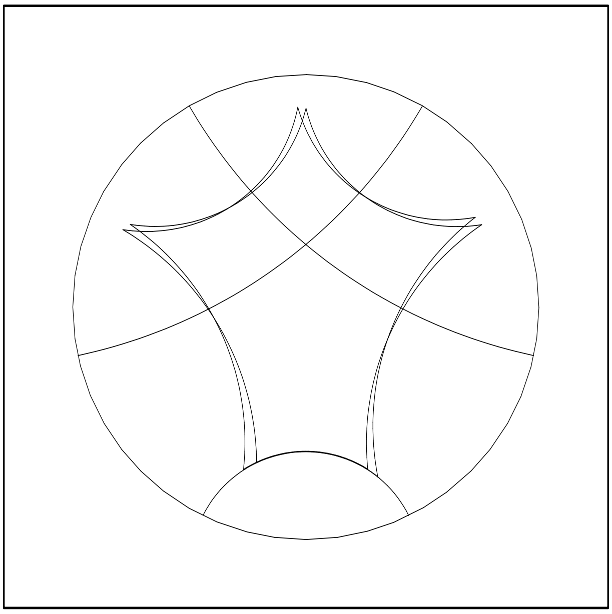

Figure 5.9: Different basepoints; non-isometric pentagons; isometric surfaces. Plotted in the disc model. The sides of the pentagons corresponding to the boundary lie on the axis of $[g^{-1}, h^{-1}]$. The other curves shown are the axes of g and h.

are fundamental domains for isometric complete hyperbolic structures on S. These pentagons are in general not isometric. See figures 5.9 and 5.10.

Now suppose p is chosen to lie slightly inside the convex core. Then $\mathcal{P}(g, h, p)$ is a fundamental domain for a submanifold of S_0, which is obtained by truncating the hyperbolic punctured torus with totally geodesic boundary along a geodesic loop parallel to the boundary. It is a hyperbolic cone-manifold with cone angle equal to the sum of the angles in the pentagon, which is greater than π. (This is clear intuitively; we will prove it shortly.) If we choose p, p' at the same distance from $\text{Axis}[g^{-1}, h^{-1}]$, then in general we obtain non-isometric pentagons corresponding to non-isometric hyperbolic cone-manifolds. The hyperbolic surface with totally geodesic boundary has been truncated along geodesic arcs based at distinct points p, p' the same distance from the boundary. The cone angle is the same in both cone-manifolds obtained: as we will see, this is because $\text{Tw}([g^{-1}, h^{-1}], p) = \text{Tw}([g^{-1}, h^{-1}], p')$. See figures 5.11, 5.12.

If instead we choose p outside the convex core of ρ, then we obtain a hyperbolic cone-manifold which can be thought of as an extension of S_0. Indeed the quotient of the entire hyperbolic plane by the discrete group of isometries determined by ρ is metrically a punctured torus which flares out past the geodesic boundary. The situation is depicted in figures 5.13, 5.14.

Now the developing map for S_0 can be seen as a tessellation by copies of $\mathcal{P}(g, h, p)$

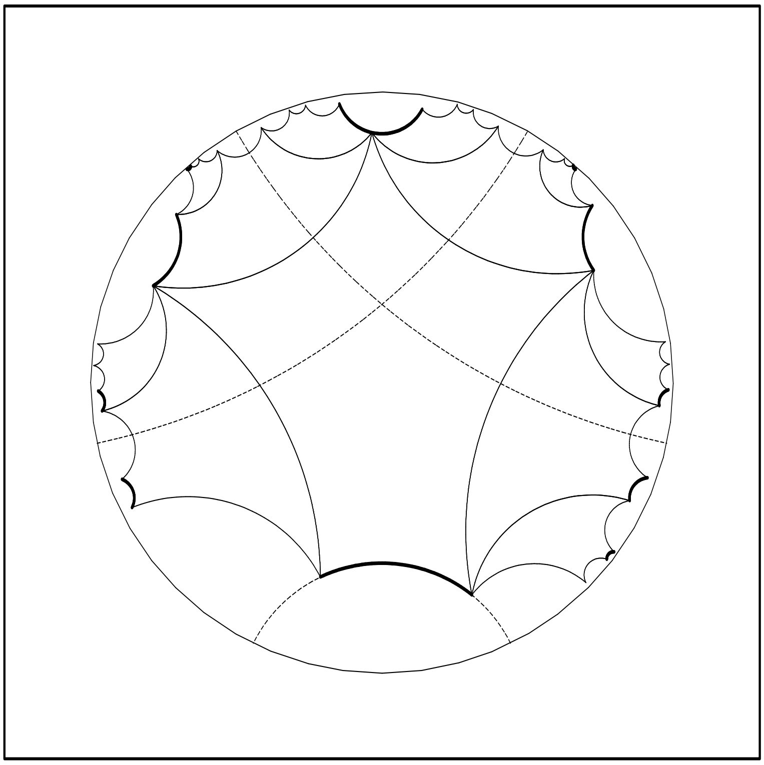

Figure 5.10: Part of the developing map. The “central” pentagon here is identical to one of the two appearing in the previous figure. Boundary edges are thickened. Axes of $g, h, [g^{-1}, h^{-1}]$ are dashed.

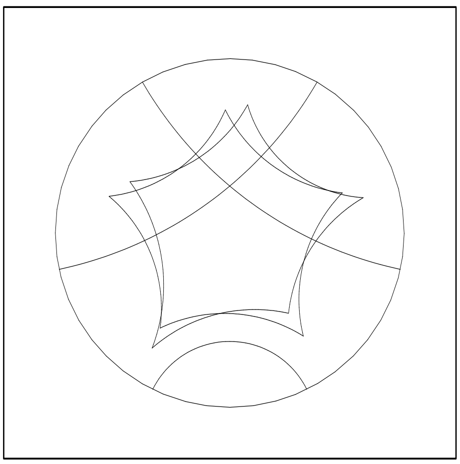

Figure 5.11: Distinct basepoints at the same distance from the axis, corresponding to a complete hyperbolic structure truncated at different points the same distance from the boundary.

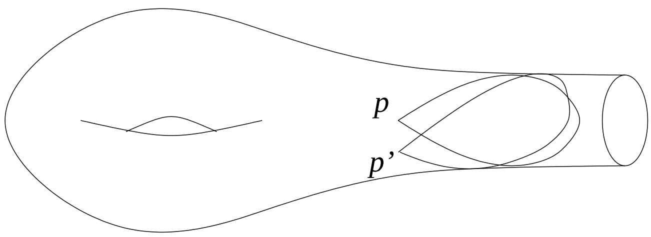

Figure 5.12: Truncation at p, p' the same distance from the boundary, creating corner points with the same angle.

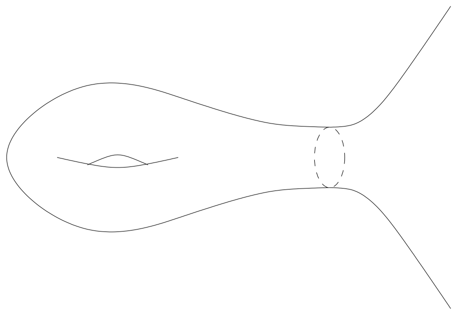

Figure 5.13: The underlying flared surface and cone-manifold obtained by taking p outside the convex core of ρ

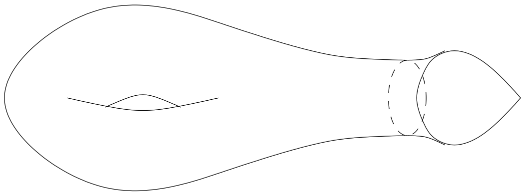

Figure 5.14: The hyperbolic cone-manifold obtained by extending beyond the geodesic boundary; equivalently, taking p outside the convex core of ρ. The corner angle is less than π.

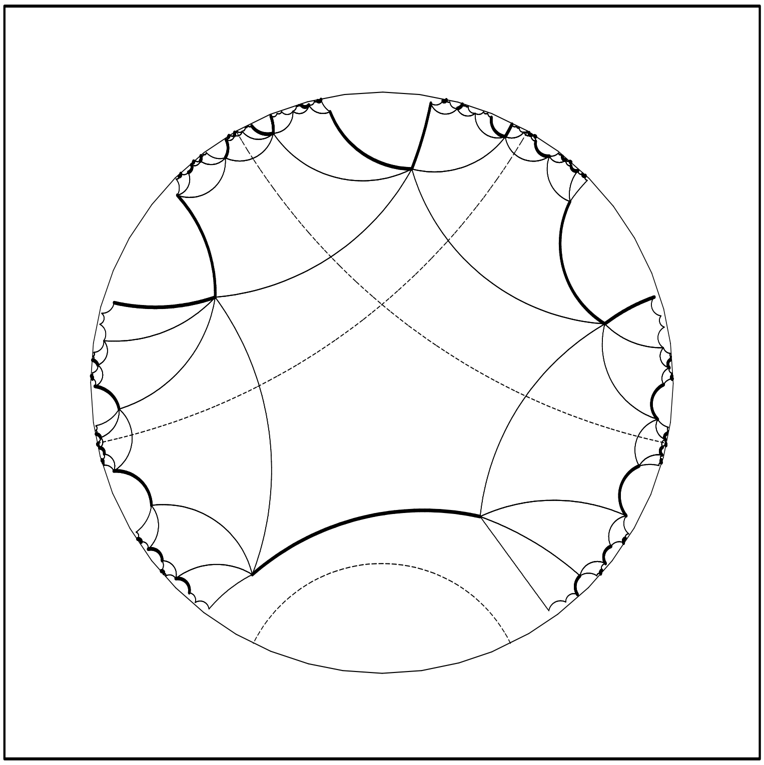

Figure 5.15: Part of a developing map for a hyperbolic cone-manifold structure obtained by truncating S_0. Again axes of $g, h, [g^{-1}, h^{-1}]$ are dashed.

filling the convex core of ρ. Other pentagons of this tessellation are of the form $\mathcal{P}(\alpha g \alpha^{-1}, \alpha h \alpha^{-1}, \alpha(p))$ where α is an element of the holonomy group, i.e. α lies in the image of ρ. See figure 5.10. Such pentagons are isometric. Similarly, a hyperbolic cone-manifold structure on S obtained by truncating or extending S_0 can be seen as a tessellation of some subset of $\mathbb{H}^2$ by non-overlapping pentagons of the form $\mathcal{P}(\alpha g \alpha^{-1}, \alpha h \alpha^{-1}, \alpha(p))$. See figures 5.15, 5.16.

It's clear that we may extend our hyperbolic cone-manifold arbitrarily far outwards from the complete structure S_0, since there is a well-defined underlying "flared" surface, depicted in figure 5.13. Thus it is possible to obtain a corner angle arbitrarily close to 0. In the other direction, it is possible to truncate S_0 with a geodesic loop based arbitrarily far from ∂S_0, but we must choose our basepoint judiciously. For instance it is possible to choose such a sequence of basepoints in $\mathbb{H}^2$ converging to the point at infinity which is an endpoint of Axis(h), as shown in figure 5.17. This gives a corner angle arbitrarily close to 2π; see figure 5.18. But for a general choice of basepoint, the geodesic loop will not be simple, and the pentagonal fundamental domain will no longer bound an immersed disc. See figure 5.19.

In the surface S_0 with totally geodesic boundary, there is a fixed width w, depending only on the length of the boundary (*not* on ρ, other than the trace of the

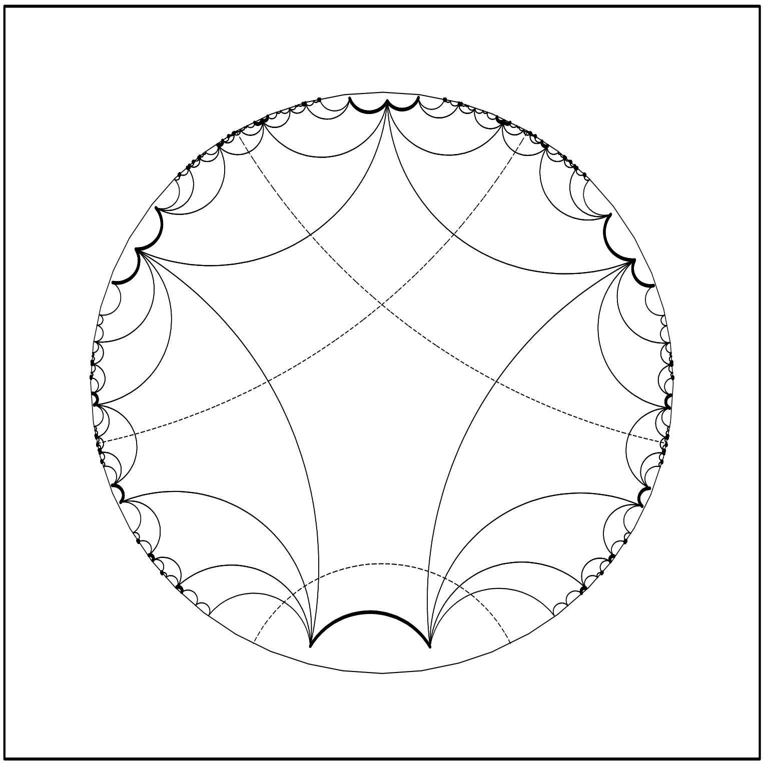

Figure 5.16: Part of a developing map for a hyperbolic cone-manifold structure obtained by extending S_0.

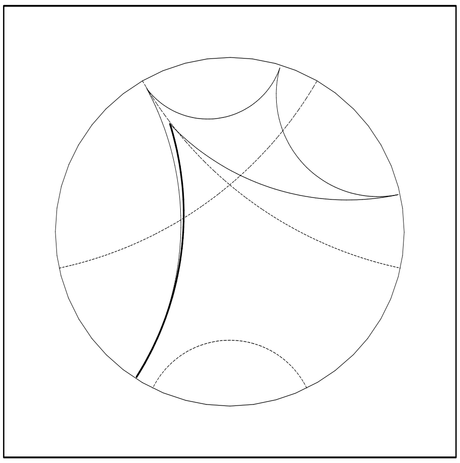

Figure 5.17: A well-chosen basepoint far inside S_0.

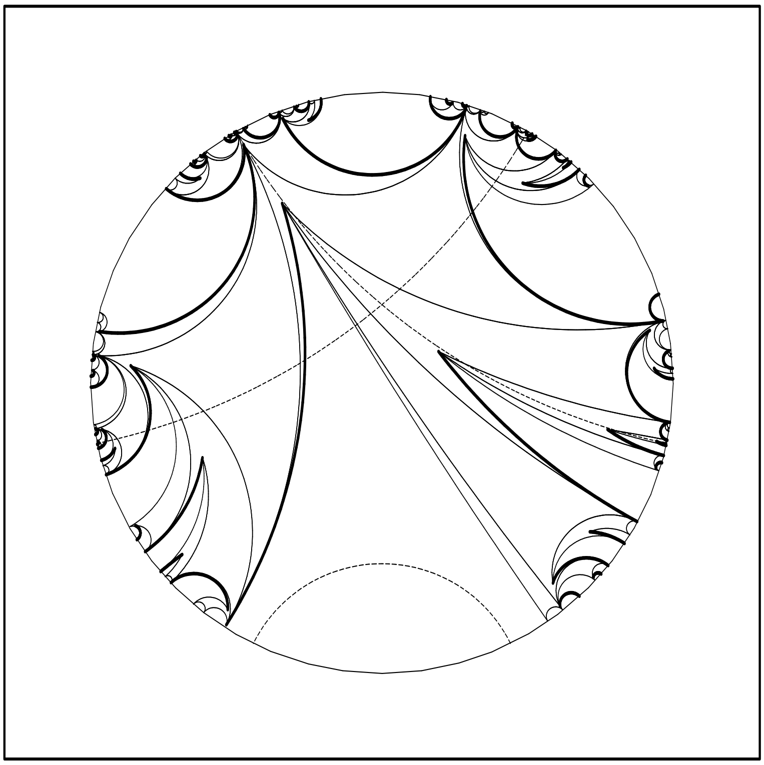

Figure 5.18: A corresponding partial developing map, with corner angle close to 2π.

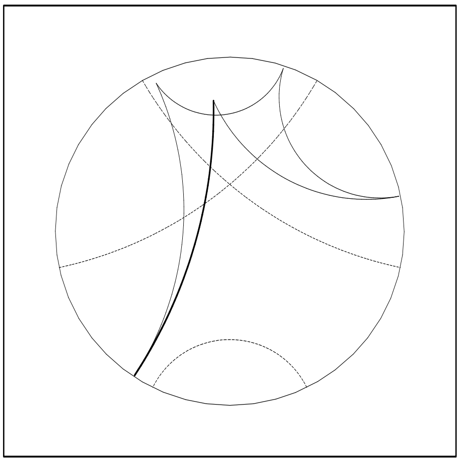

Figure 5.19: Choosing the basepoint injudiciously gives a non-simple pentagon; the geodesic arc in S_0 of truncation is non-simple.

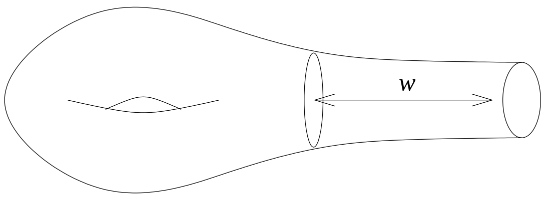

Figure 5.20: The collar on S_0

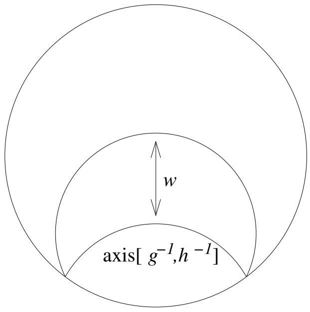

Figure 5.21: Developing map of the collar on S_0

image of the boundary curve), called the *collar width*, such that the set of all points within w of the boundary forms a topological cylinder called the *collar*. (See e.g. [10] for details.) We may truncate the punctured torus at any point inside the collar: see figure 5.20. The boundary of the collar consists topologically of two circles, namely ∂S_0 and a constant distance curve at distance w from ∂S_0. The collar region develops to a region in the convex core of ρ consisting of points at distance $\leq w$ from Axis$[g^{-1}, h^{-1}]$: see figure 5.21. Such a region is convex, and hence there is a simple geodesic arc in the collar region connecting any two points in the region. Since we are truncating a complete hyperbolic surface S_0, any choice of p in the convex core of ρ, within a distance of w from Axis$[g^{-1}, h^{-1}]$, gives $\mathcal{P}(g, h, p)$ simple, bounding an embedded disc. That is, our choice of p need not be judicious in this region!

The many hyperbolic cone-manifold structures considered in the preceding pages have come from one particular type of representation ρ: a representation which is the holonomy of a hyperbolic structure on S with totally geodesic boundary. (Indeed all the plots in the hyperbolic plane have been generated from the same two isometries g, h!) As we will see these are precisely the representations with $\mathrm{Tr}[g, h] < -2$.

When $\mathrm{Tr}[g, h] = -2$ we will see that ρ is discrete also: ρ in this case is the holonomy of a complete structure on a cusped torus. This is a limiting case of the complete structure on the punctured torus with totally geodesic boundary described above. If we choose p to be the fixed point at infinity of $[g^{-1}, h^{-1}]$ then $\mathcal{P}(g, h, p)$ degenerates to a quadrilateral, and we obtain precisely this cusped torus. Different choices for p correspond to truncating this cusped torus, and as we choose p further

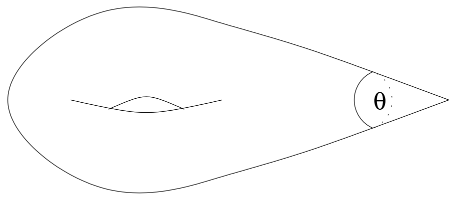

Figure 5.22: A "pinched" torus, i.e. with one cone point, angle in $(0, 2\pi)$. We have $\text{Tr}[g, h] \in (-2, 2)$ in this case.

from the circle at infinity, the corner angle θ increases.

However, there are representations ρ which are not discrete but which are the holonomy of a cone-manifold structure on the punctured torus with no cone points and one corner point. In fact we will see that every representation which is not virtually abelian is a holonomy representation of this type. When ρ is not discrete, we cannot think of a cone-manifold structure on S as a complete structure which has been truncated or extended. Rather the punctured torus is, in some sense, inherently badly formed so that it cannot be so extended. In this case also, ρ is the holonomy of many non-isometric structures on S, obtained by perturbing the basepoint p of $\mathcal{P}(g, h, p)$.

Still fixing a choice of basis curves, let us consider a representation ρ with $-2 < \text{Tr}[g, h] < 2$, i.e. where the boundary of S is taken to an elliptic element. Such a representation ρ is the holonomy of a hyperbolic cone-manifold structure on a "punctured" torus with a single corner point and a boundary of length 0! This is better thought of as a (non-punctured) torus T with a single (interior) cone point, with angle in $(0, 2\pi)$ (see figure 5.22). Indeed a fundamental domain for the developing map is $\mathcal{P}(g, h, p)$, where p is the fixed point of the elliptic isometry $[g^{-1}, h^{-1}]$; the pentagon degenerates to a quadrilateral bounding an embedded disc (see figure 5.23). However as the representation is not discrete, the developing map is in general not a nice tessellation of these quadrilaterals, but overlaps itself. See figures 5.24 and 5.28.

By perturbing the point $p \in \mathbb{H}^2$ in certain directions, we obtain a hyperbolic cone-manifold structure on S with no cone points and one corner point (see figures 5.25 and 5.26). The corner angle might lie anywhere in $(0, 3\pi)$. But there is a sense in which the representation ρ is inherently "large angle" or "small angle". For certain ρ, perturbing p can only produce corner angles in $(0, 2\pi)$; for other ρ, perturbing p can only produce corner angles in $(2\pi, 3\pi)$. The reason for this, we will see, is essentially because $[g, h]$ is well-defined in $\widetilde{PSL_2\mathbb{R}}$ and hence determines a rotation of an angle specified as a real number rather than modulo 2π. In the "small angle" case, a structure obtained on S can be thought of as a truncation of

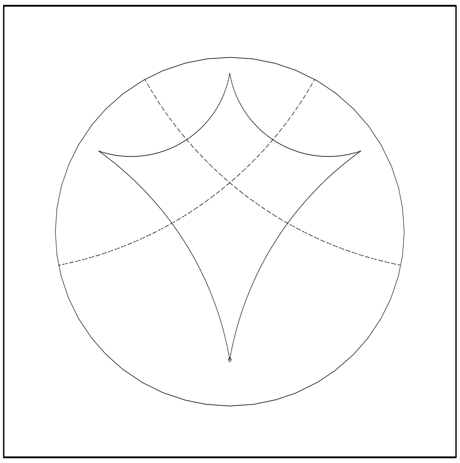

Figure 5.23: A fundamental domain for the developing map of T in the elliptic boundary case. The axes of g, h are dotted. The fixed point of $[g^{-1}, h^{-1}]$ is marked.

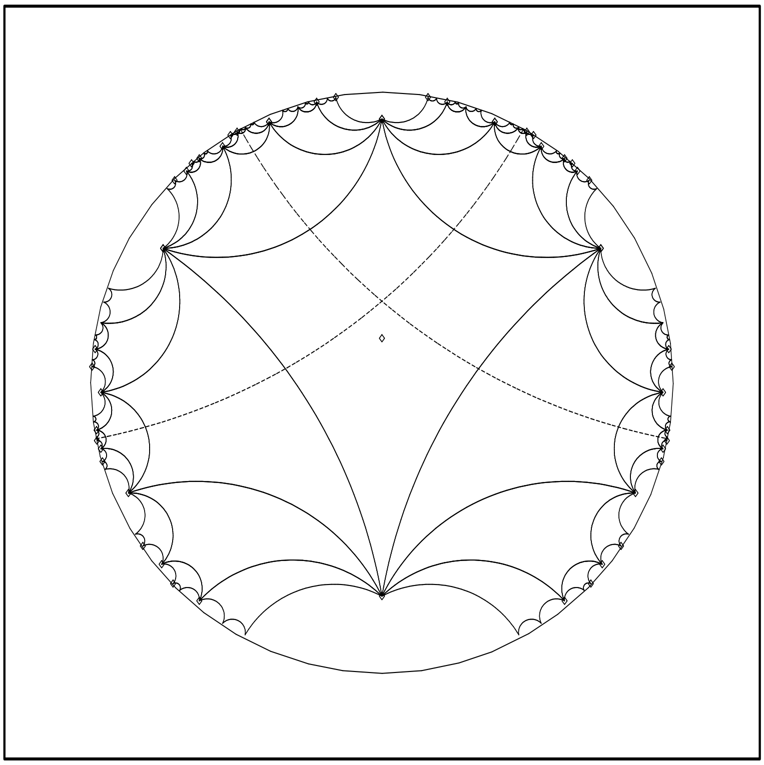

Figure 5.24: A corresponding partial developing map in the elliptic boundary case. As usual the axes of g, h are dotted.

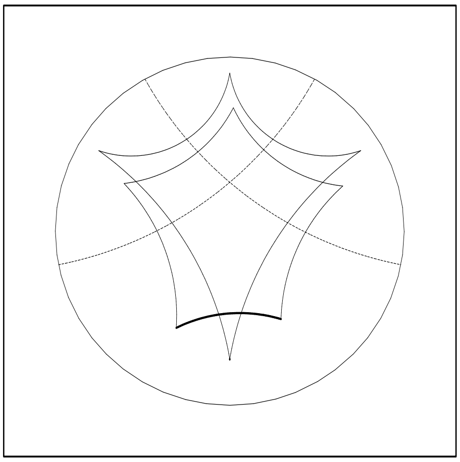

Figure 5.25: Perturbing p produces a fundamental domain for a punctured torus S, which can be regarded as a restriction of T.

the hyperbolic cone-manifold T described above. See figures 5.24, 5.25 and 5.26.

In the "large angle" case, a hyperbolic cone-manifold structure on S is obtained which perhaps be thought of as an "extension" of T. But the extension is of such a nature that the cone angle is always larger than 2π. Consider figures 5.27, 5.28, 5.29.

Note that many of the figures have the axes of g and h included in them. Recall g identifies two sides of $\mathcal{P}(g,h,p)$. If $\text{Axis}\, g$ intersects $\mathcal{P}(g,h,p)$ and in a segment and g identifies the points at the end of the segment, then on the surface S, the axis of g becomes a closed geodesic loop in the free homotopy class of G, which avoids the corner point. The same applies to h. This is just the standard "straightening" process for loops on hyperbolic manifolds (see, e.g., [60]). In all cases, with $\text{Tr}[g,h] < 2$, $\text{Axis}(g)$ and $\text{Axis}(h)$ intersect (lemma 3.2.2). Provided we do not perturb the point p too far from p_0, the fixed point of $[g^{-1}, h^{-1}]$, we may thus obtain geodesics on S in the free homotopy classes of G, H which avoid the corner point and intersect at precisely one point.

We will see in the next chapter that the case $\text{Tr}[g,h] = 2$ corresponds to reducible representations; for these we will give special constructions. There are also the representations with $\text{Tr}[g,h] > 2$. Hyperbolic cone-manifold structures obtained in these cases possess similar non-rigidity properties, and are again in a sense inherently malformed, but are uglier than the previous cases. Given such g, h it is often difficult, or impossible, to find p such that $\mathcal{P}(g,h,p)$ is a simple pentagon bounding an immersed disc. One can try to change the basis curves G, H to improve the situation.

One indication of the inherent ugliness in this case is that, in any hyperbolic cone-manifold structure obtained on S with no cone points and one corner point,

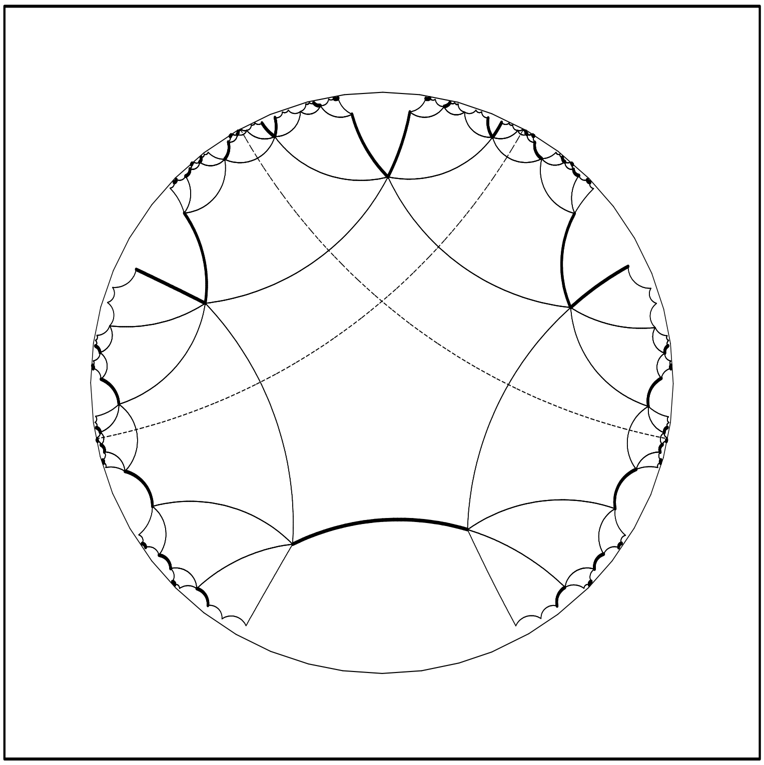

Figure 5.26: A partial developing map in the "small angle" case. Note the correspondence with the developing map on T.

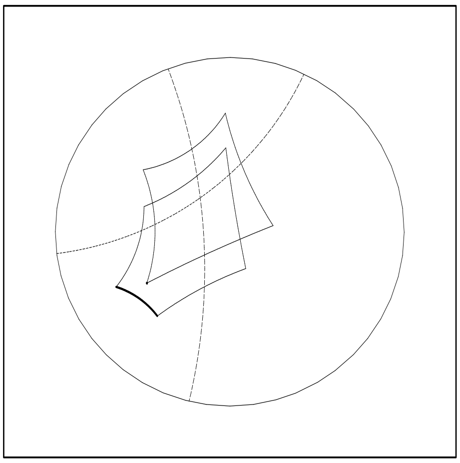

Figure 5.27: Fundamental domains for T and S in the "large angle" case.

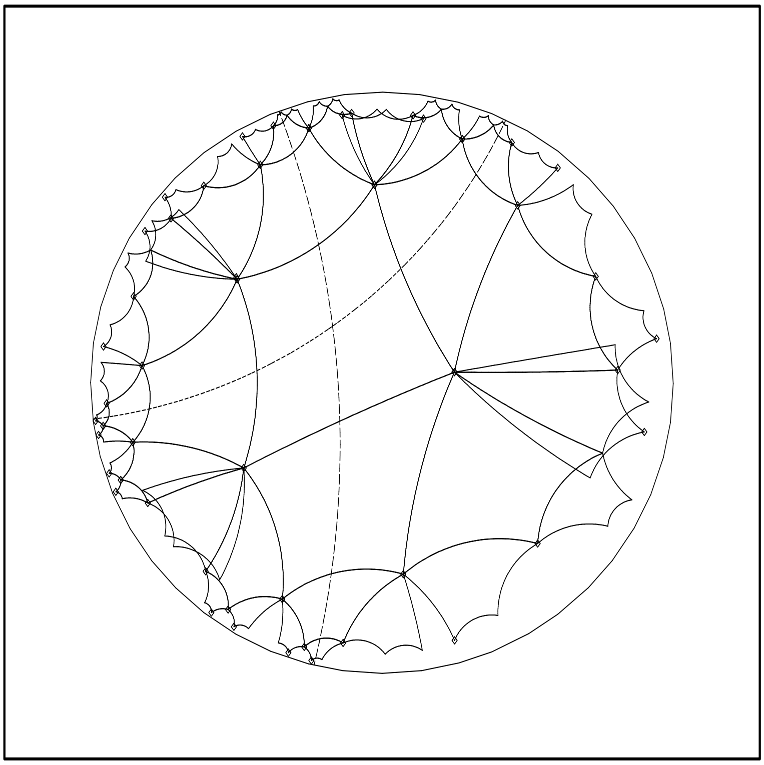

Figure 5.28: Partial developing map for T in the “large angle” case.

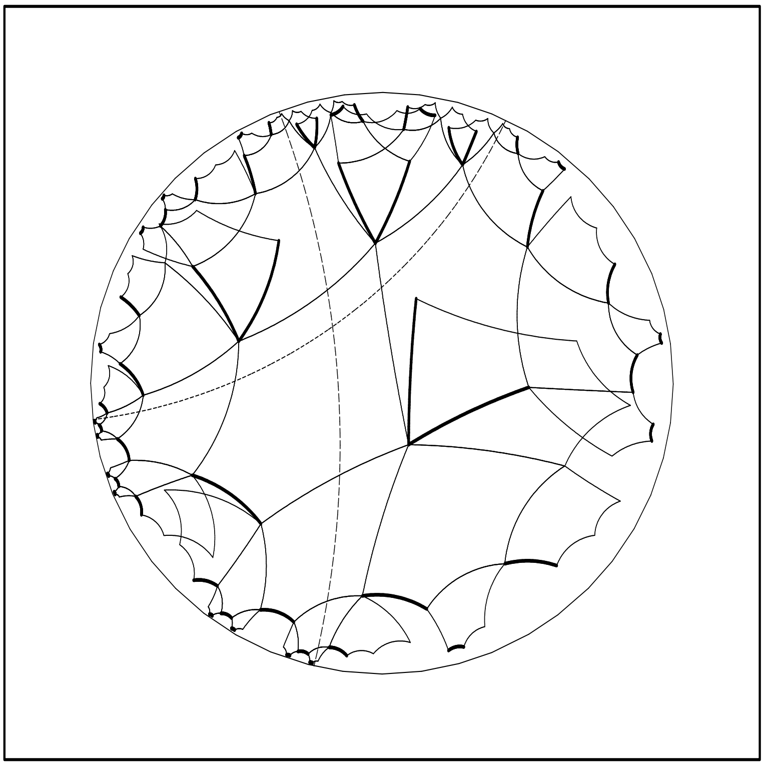

Figure 5.29: Partial developing map for S in the “large angle” case. Note the correspondence with the developing map for T.

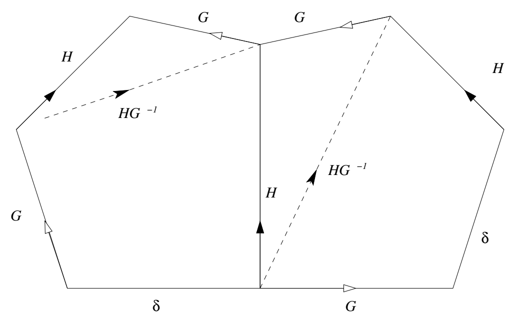

Figure 5.30: A change of basis corresponding to a simple cut and paste of pentagons.

there are no closed geodesics γ_G, γ_H on S in the free homotopy classes of G, H avoiding the corner point. Otherwise, as basis curves lying in the interior of S, γ_G and γ_H intersect. Considering a developing map $\mathcal{D}$ and pentagonal fundamental domain, g has to translate along $\mathcal{D}(\gamma_G)$, and hence would be hyperbolic. Similarly h must be hyperbolic, and the axes of g and h intersect. This contradicts $\text{Tr}[g, h] > 2$.

Thus concludes our preliminary discussion of one aspect of non-rigidity: the choice of the basepoint p. We have considered G, H to be fixed basis curves throughout our discussion. The other aspect of non-rigidity is in the choice of basis curves G, H. If we change the basis G, H of $\pi_1(S, q)$ to another basis G', H', we may obtain another pentagon $\mathcal{P}(g', h', p')$ describing a hyperbolic cone-manifold structure on S. The change of basis $(G, H) \mapsto (G, HG^{-1})$ for instance has a simple geometric interpretation as cutting our pentagon along a diagonal and regluing, arising from a Dehn twist in S. See figure 5.30.

It's a well-known result that the mapping class group of a surface is generated by Dehn twists. We will see in the next chapter that, in the case of a punctured torus, the outer automorphism group of $\pi_1(S, q)$ corresponds precisely to the mapping class group. Thus any basis change has a cut-and-paste interpretation.

However, in a non-convex pentagon a diagonal may not lie entirely inside, so that even from the same basepoint the pentagon $\mathcal{P}(g', h', p)$ obtained by change of basis may no longer bound an immersed disc. So with injudicious changes of basis we may lose our pentagonal fundamental domain of the hyperbolic cone-manifold structure on S. But, as we will see later, judicious changes of basis may allow us to demonstrate an explicit hyperbolic cone-manifold structure.

5.3 Geometric parameters

In this section we will search for geometric quantities which uniquely determine a cone-manifold structure on S with no interior cone points and at most one corner

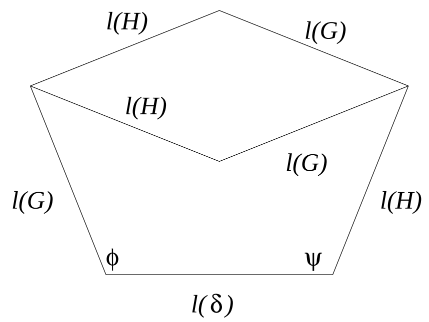

Figure 5.31: A pentagon is almost determined by its lengths and 2 angles.

point. It is clear that a pentagon $\mathcal{P}(g, h, p)$ determines the structure on S completely; but g, h are rather algebraic. The three lengths of sides $l(\partial), l(G), l(H)$ combined with the five angles of the pentagon are geometric quantities which specify the structure completely, but there is some redundancy.

The three lengths $l(\partial), l(G), l(H)$, and the two angles ϕ, ψ of the pentagon adjacent to the edge corresponding to the boundary of S define the pentagon *almost* uniquely. Assuming the pentagon to be of a specified orientation, there are in general two possibilities, defined by the intersection of two circles of radii $l(G), l(H)$, as shown in figure 5.31.

Thus one other piece of information is necessary: say the total cone angle θ. A pentagon is defined uniquely by the six parameters $(l(\partial), l(G), l(H), \phi, \psi, \theta)$, which are all part of the geometric data from S.

However, as discussed above there are non-isometric pentagons giving rise to isometric cone-manifold structures on S. Hence there are distinct sextuples giving rise to isometric cone-manifold structures on S. We would like to have a canonical sextuple describing a hyperbolic cone-manifold structure on S up to isometry.

The procedure given in section 5.1 above almost allows us to do this; the procedure is almost (but not quite) canonical. A canonical pentagonal decomposition would give us a canonical set of six parameters for the hyperbolic structure. In general however, if there are several non-homotopic shortest length geodesics then we have choices to make. Nevertheless, it is possible to make some simple (perhaps not aesthetically satisfactory) alterations to make this procedure canonical. For instance, amongst all shortest non-boundary-parallel curves G_i through the basepoint p (there are only finitely many), choose the one for which the associated H_i is shortest; if there are still several choices then choose the one such that ϕ is least.

In the cases where $\text{Tr}[g, h] \in (-\infty, -2), (-2, 2)$ discussed above, where the structure on S can be considered as a truncation or extension of a structure with totally geodesic boundary, or a "perturbation" of a complete structure on the torus with a cone point, we may consider the "natural" non-truncated or non-perturbed surface associated and take shortest meridian and longitude curves G_0, H_0 in the free homotopy class of G, H as above. Then we may take $l(G_0), l(H_0), l(\partial)$ as our parameters, which specify the underlying "natural" surface (i.e. complete hyperbolic punctured torus with totally geodesic boundary, or hyperbolic torus with a cone point) uniquely. For a complete hyperbolic structure with totally geodesic boundary, this is a well-known fact, say as shown in [60]. In any case it follows from facts proved in the next chapter that knowing $\mathrm{Tr}\, g, \mathrm{Tr}\, h, \mathrm{Tr}[g,h]$, we can determine $\mathrm{Tr}\, gh$ by solving a quadratic; (for an irreducible representation) knowing $\mathrm{Tr}\, g, \mathrm{Tr}\, h, \mathrm{Tr}\, gh$ determines the representation uniquely up to conjugacy; thus the underlying surface is determined uniquely; and the two roots of the quadratic correspond to the same representation after a change of basis (corresponding to to a Markoff move as discussed in the next chapter), so give the same underlying surface. Once the underlying surface is specified, then two more parameters describe where we place the corner point q: say θ (which determines the distance from the boundary, as we will see in the next chapter) and φ describing where along the constant distance curve q is placed. Thus we may take $(l(G), l(H), l(\partial), \theta, \varphi)$ defines our surface uniquely. The cases $\mathrm{Tr}[g,h] > 2$ are less easily given to geometry of this sort, as we shall see in the construction of geometric structures in chapter 7.

5.4 Twisting and the corner angle

We will now investigate further the developing map of a hyperbolic cone-manifold structure on a punctured torus, in particular the action of $\pi_1(S)$ as isometries of $\mathbb{H}^2$ via the holonomy representation, and their derivatives. We will find a simple relationship between the twisting involved and the corner angle obtained.

The pentagon $\mathcal{P}(g,h,p)$ is a fundamental domain and isometric copies of $\mathcal{P}$, possibly overlapping, give the entire developing map $\mathcal{D} : \tilde{S} \longrightarrow \mathbb{H}^2$. Recall that g and h identify pairs of sides of $\mathcal{P}$, and the vertices of $\mathcal{P}$ are images of p under various combinations of g and h. Notate the vertices as

$$p_0 = p,\ p_1 = hp,\ p_2 = ghp,\ p_3 = h^{-1}ghp,\ p_4 = g^{-1}h^{-1}ghp.$$

Let the corresponding angles of the pentagon be $\theta_0, \ldots, \theta_4$, so that their sum is equal to the corner angle θ. We now have the following relationship.

Lemma 5.4.1 *If the segment $p \to [g^{-1}, h^{-1}]p$ bounds $\mathcal{P}(g,h,p)$ on its left then $\theta \equiv \pi - \mathrm{Tw}([g^{-1}, h^{-1}], p)$ modulo 2π. If the segment $p \to [g^{-1}, h^{-1}]p$ bounds $\mathcal{P}(g,h,p)$ on its right then $\theta \equiv \pi + \mathrm{Tw}([g^{-1}, h^{-1}], p)$ modulo 2π.*

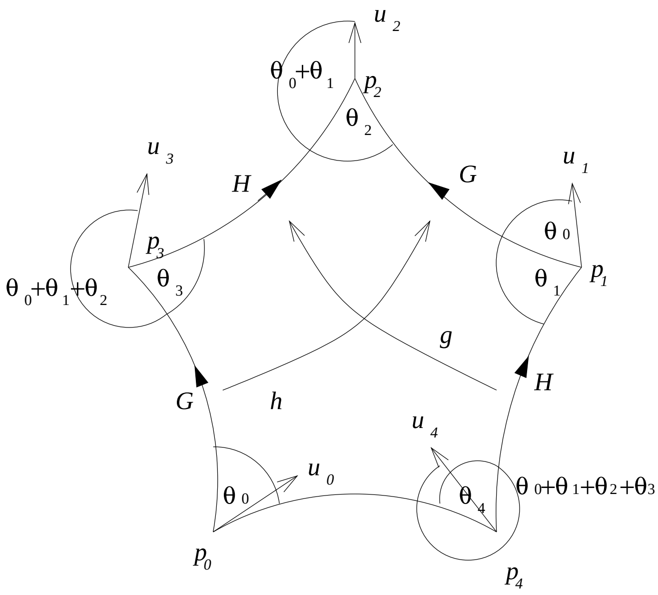

Figure 5.32: A unit vector chase.

Proof First we consider the orientation shown in figure 5.32. Consider a unit tangent vector (p, u) at p pointing along the geodesic to $p_4 = [g^{-1}, h^{-1}]p$, i.e. the positive direction of the boundary curve, and let the curves G, H be oriented as shown. Follow the image of this vector under Dh, Dg, Dh^{-1} and Dg^{-1} to obtain unit tangent vectors (p_i, u_i) at each p_i.

Now $(p_0, u_0) = (p, u)$ is based at p and points θ_0 clockwise of the positive direction of G; hence $(p_1, u_1) = Dh(p, u)$ is based at p_1 and points θ_0 clockwise of the positive direction of G. But (p_1, u_1) points $\theta_0 + \theta_1$ clockwise of the negative direction of H, hence so does $(p_2, u_2) = Dg(p_1, u_1)$. Then (p_2, u_2) lies $\theta_0 + \theta_1 + \theta_2$ clockwise of the negative direction of G, hence so does $(p_3, u_3) = Dh^{-1}(p_2, u_2)$. Finally, (p_3, u_3) lies $\theta_0 + \theta_1 + \theta_2 + \theta_3$ clockwise of the positive direction of H, and so does $(p_4, u_4) = D[g^{-1}, h^{-1}](p, u)$. Thus $D[g^{-1}, h^{-1}](p, u)$ lies $\theta_0 + \theta_1 + \theta_2 + \theta_3 + \theta_4 = \theta$ clockwise of the boundary segment. Then we see that $\mathrm{Tw}([g^{-1}, h^{-1}], p) = \pi - \theta$, modulo 2π.

If $\mathcal{P}(g, h, p)$ has the other orientation, where $p \to [g^{-1}, h^{-1}]p$ bounds $\mathcal{P}(g, h, p)$ on its right, then we obtain similarly $\mathrm{Tw}([g^{-1}, h^{-1}], p) = -\pi + \theta$, modulo 2π. ■

Chapter 6

The Algebra of Punctured Tori

We will shortly be finding geometric structures on the punctured torus with given holonomy. Consequently it will be important to analyse representations of the free group on two generators into the group $PSL_2\mathbb{R}$. Throughout this chapter, S denotes a punctured torus, simply as a topological object. We do not yet consider geometric structures. We take a basepoint q on ∂S. We write $\pi_1(S)$ for the fundamental group, with the basepoint q implied.

6.1 Characters of representations

Let G, H be a basis of $\pi_1(S)$, so $\pi_1(S) = \langle G, H \rangle$. A representation $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ or $SL_2\mathbb{R}$ is determined by $\rho(G)$ and $\rho(H)$. A representation into $PSL_2\mathbb{R}$ obviously lifts to $SL_2\mathbb{R}$, and we have two choices each for the lifts of $\rho(G)$ and $\rho(H)$. For now consider ρ as a representation into $SL_2\mathbb{R}$ and denote $\rho(G) = g$, $\rho(H) = h$.

From our point of view, conjugate holonomy representations are equivalent, and give isometric collections of geometric structures (although, as seen in the previous chapter, a single representation can give many geometric structures). So we are only interested in conjugacy classes and characters of representations, as discussed in section 4.4. Recall from there that the character of ρ is determined by the value of $\text{Tr} \circ \rho$ at finitely many elements of $\pi_1(S)$. For a punctured torus with $\pi_1(S) = \langle G, H \rangle$, it is sufficient to consider only the three elements G, H, GH. For any word W in G, H and their inverses, we can write $\text{Tr}(\rho(W))$ as a polynomial in

$$(x, y, z) = (\text{Tr}\, g,\ \text{Tr}\, h,\ \text{Tr}\, gh).$$

For instance we can obtain the important relation

$$\text{Tr}[g, h] = \text{Tr}^2 g + \text{Tr}^2 h + \text{Tr}^2 gh - \text{Tr}\, g \, \text{Tr}\, h \, \text{Tr}\, gh - 2$$

and hence we define the polynomial

$$\kappa(x, y, z) = x^2 + y^2 + z^2 - xyz - 2.$$

following notation of [28], [30]. For more details see [43], [15], [30], [28, 4.1], [41, 3.4] or [22].

It is a classical result that for *irreducible* ρ, the triple $(\mathrm{Tr}\, g, \mathrm{Tr}\, h, \mathrm{Tr}\, gh)$ not only defines the character of ρ, but actually defines the pair $g, h \in SL_2\mathbb{R}$ uniquely up to conjugacy: see [28], [21], [22] for a proof. Recall a representation into $SL_2\mathbb{R}$ is *reducible* if its image is a set of matrices which, acting via linear transformations on $\mathbb{C}^2$, leaves invariant a line in $\mathbb{C}^2$.

Theorem 6.1.1 *Let G, H be a free basis of $\pi_1(S)$ and suppose $\rho : \pi_1 \longrightarrow SL_2\mathbb{R}$ is irreducible. Then the traces of g, h, gh determine ρ up to conjugacy.* ■

The set of all $(x, y, z) = (\mathrm{Tr}\, g, \mathrm{Tr}\, h, \mathrm{Tr}\, gh) \in \mathbb{R}^3$ is the character variety $X(S)$ of S. It is not all of $\mathbb{R}^3$, and the following theorem describes $X(S)$ exactly. We refer to [28, thm. 4.3] for a proof.

Theorem 6.1.2 *For given $(x, y, z) \in \mathbb{R}^3$, there exist $g, h \in SL_2\mathbb{R}$ such that*

$$(\mathrm{Tr}\, g, \mathrm{Tr}\, h, \mathrm{Tr}\, gh) = (x, y, z)$$

if and only if

(i) $\mathrm{Tr}[g, h] = x^2 + y^2 + z^2 - xyz - 2 \geq 2$*; or*

(ii) at least one of $|x|, |y|, |z|$ is ≥ 2. ■

(Actually if we have $\mathrm{Tr}[g, h] < 2$ then $\mathrm{Tr}[g, gh] < 2$ also, so by lemma 3.2.2 all of g, h, gh give hyperbolic isometries of $\mathbb{H}^2$ and hence *all* $|x|, |y|, |z| > 2$.) Thus the set of $(x, y, z) \in \mathbb{R}^3$ without corresponding representations (i.e $\mathbb{R}^3 \backslash X(S)$) are those with $\kappa(x, y, z) < 2$ and $-2 < x, y, z < 2$: see figure 6.1.

For representations $\pi_1(S) \longrightarrow PSL_2\mathbb{R}$, the character variety can be described simply also. There are four ways to lift $\rho(G), \rho(H)$ into $SL_2\mathbb{R}$, which are related by sign changes. Thus we simply take the character variety $X(S)$ of representations into $SL_2\mathbb{R}$ modulo the equivalence relation

$$(x, y, z) \sim (-x, -y, z) \sim (-x, y, -z) \sim (x, -y, -z)$$

induced by these four possible lifts. The notion of reducibility still makes sense: elements of $PSL_2\mathbb{R}$ act via linear transformations on $\mathbb{C}^2$ up to a reflection in the origin, hence on $\mathbb{C}P^1$, so the idea of an invariant line still makes sense. And for

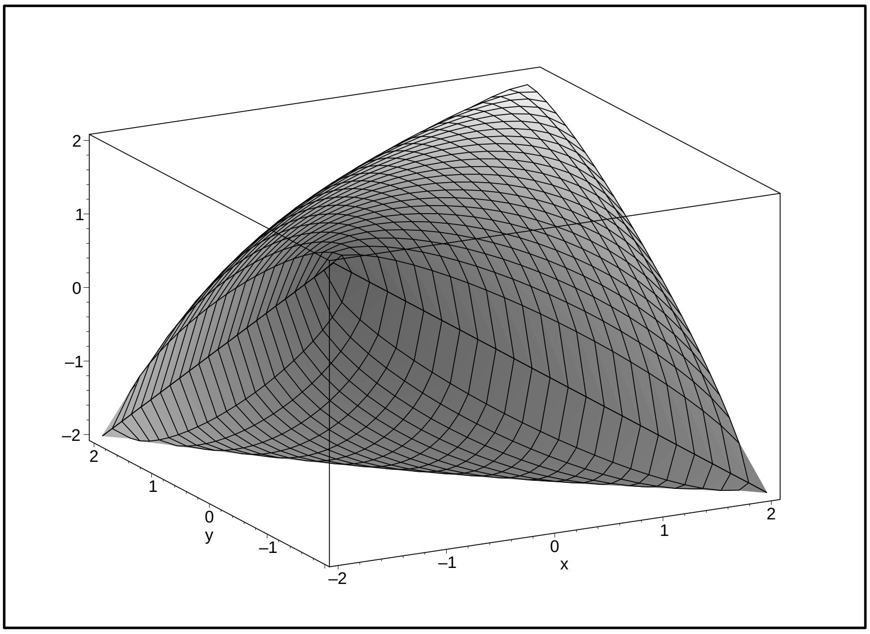

Figure 6.1: $\mathbb{R}^3 \backslash X(S)$, i.e. the region of $\mathbb{R}^3$ without representations: (strictly) inside this curved tetrahedron-like surface.

representations into $PSL_2\mathbb{R}$ the value of $\kappa(x, y, z) = \text{Tr}[g, h]$ is well-defined, even if the signs of x, y, z are ambiguous.

The reducible representations have a simple characterisation: see [15] or [30] for a proof.

Proposition 6.1.3 *The representation $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ or $SL_2\mathbb{R}$ is reducible if and only if* $\text{Tr}[g, h] = 2$*, i.e. iff the character (x, y, z) of ρ satisfies $\kappa(x, y, z) = 2$.* ■

Note this trivially implies that an abelian representation is reducible: for representations into $SL_2\mathbb{R}$ we see $[g, h] = I$ so $\text{Tr}[g, h] = 2$. For representations into $PSL_2\mathbb{R}$ we see $[g, h] = \pm I$; but by the classification in corollary 3.7.3 this implies $\text{Tr}[g, h] = 2$ also.

We have now defined the character variety $X(S) \subset \mathbb{R}^3$. Points with $\kappa(x, y, z) = 2$ describe reducible representations, which include abelian representations. Points with $\kappa(x, y, z) \neq 2$ describe irreducible representations, hence describe a conjugacy class of representations precisely. For $t \neq 2$, the space of all representations (up to conjugacy) with $\text{Tr}[g, h] = t$ can be identified with the set $X_t(S) = \kappa^{-1}(t) \cap X(S)$: this is a relative character variety of S as described in section 4.4.

In principle, for irreducible ρ it is possible to deduce all the geometry of g and h, considered as isometries of the hyperbolic plane, from the triple $(\text{Tr}\, g, \text{Tr}\, h, \text{Tr}\, gh)$. This partially motivates results such as those proved in section 3.2.

6.2 Nielsen's Theorem

So far we have only considered one particular basis G, H of $\pi_1(S)$. There are many different bases and any two bases (G, H) and (G', H') are related by an automorphism. It is a remarkable fact that, in a certain sense, 'some elements are more arbitrary than others'. We have the following theorem due to Nielsen. See [51], [42, thm. 3.9] or [40, prop. 5.1]. The proof relies upon Nielsen's description of the automorphism group of the free group on two generators.

Theorem 6.2.1 (Nielsen) *An automorphism ϕ of $\langle G, H\rangle$ takes $[G, H]$ to a conjugate of itself or its inverse $[H, G]$.* ∎

We call the automorphism ϕ either *orientation-preserving* or *orientation-reversing* as $[G, H]$ is taken respectively to a conjugate of itself or of $[H, G]$. We can therefore say whether two bases (G, H), (G', H') have the same or different orientation.

Thus the conjugacy class of the commutator $[G, H]$ is in some sense "canonical". It is therefore natural to classify representations $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ or $SL_2\mathbb{R}$ via the trace of the commutator $[g, h]$.

We saw in section 4.5 the Dehn–Nielsen theorem: for a closed surface, every automorphism of the fundamental group is induced by a homeomorphism of the surface. This is also true for the punctured torus (see e.g. [30]). From a punctured torus S with $\pi_1(S) = \langle G, H\rangle$ we consider its double, i.e. a genus 2 surface S_2, and we can take a fundamental group presentation $\pi_1(S_2) = \langle G_1, H_1, G_2, H_2 \mid [G_1, H_1] = [G_2, H_2]\rangle$. Given an automorphism ϕ of $\pi_1(S)$ we obtain an automorphism ϕ_2 of $\pi_1(S_2)$ by setting $\phi_2(G_1)$ to be the same word in G_1, H_1 as $\phi(G)$ is in G, H. Similarly we set $\phi_2(G_2), \phi_2(H_1), \phi_2(H_2)$. This is clearly an automorphism of $\pi_1(S_2)$. By the Dehn-Nielsen theorem we obtain a homeomorphism $f : S_2 \longrightarrow S_2$. Let C denote the curve on S_2 separating our two copies of S. Since $\phi([G_1, H_1])$ is conjugate to $[G_1, H_1]^{\pm 1}$, $f(C)$ is freely homotopic to C. So f is homotopic to a homeomorphism of S_2 preserving C. By construction of ϕ_2, $\phi_2(G_1), \phi_2(H_1)$ are words in G_1, H_1, so f preserves each side of C. In particular we obtain a homeomorphism of S which induces the automorphism ϕ.

A similar result is not true for any other hyperbolic surface with nonempty boundary. See [30], and for further details see [57] and [52]. We record this result.

Theorem 6.2.2 *Any automorphism ϕ of $\pi_1(S) = \langle G, H\rangle$ is induced from a homeomorphism of S.* ∎

Consequently, as for closed surfaces, we obtain $\mathrm{MCG}(S) \cong \mathrm{Out}\,\pi_1(S)$.

6.3 The action on the character variety

Now we consider the effect of changing basis $(G, H) \mapsto (G', H')$ on a representation $\rho : \pi_1(S) \longrightarrow SL_2\mathbb{R}$. The group $\mathrm{Aut}\,\pi_1(S)$ acts simply and transitively on bases (G, H) of $\pi_1(S)$, so this is equivalent to considering the effect of an automorphism on the representation ρ, as discussed in section 4.5. Changing basis merely changes a pair of curves on S, which are quite arbitrary; but the character may change:

$$(x, y, z) = (\mathrm{Tr}\, g, \mathrm{Tr}\, h, \mathrm{Tr}\, gh) \mapsto (\mathrm{Tr}\, g', \mathrm{Tr}\, h', \mathrm{Tr}\, g'h') = (x', y', z').$$

We obtain an action of $\mathrm{Aut}\,\pi_1(S)$ on $X(S)$, which descends to an action of $\mathrm{Out}\,\pi_1(S) \cong \mathrm{MCG}(S)$. Points in $X(S)$ which are related under this action ought to be considered as equivalent in terms of the underlying geometry. We will find which elements of $X(S)$ are equivalent under this action.

Before we embark on this task however, we make some observations. In the situation above, by Nielsen's theorem, $[G, H]$ is conjugate to $[G', H']^{\pm 1}$, so $\mathrm{Tr}[g, h] = \mathrm{Tr}[g', h']$ and we have $\kappa(x, y, z) = \kappa(x', y', z')$. That is, (x, y, z) and (x', y', z') lie on the same level set of the polynomial κ. Therefore, the action of $\mathrm{Out}\,\pi_1(S)$ on the character variety $X(S)$ preserves each relative character variety $X_t(S) = \kappa^{-1}(t) \cap X(S)$. Since each $X_t(S)$ is 2-dimensional, the 2-form ω described in section 4.4 defines a symplectic structure on each $X_t(S)$, which is an area form. Recall from section 4.5 that ω is invariant under the action of $\mathrm{Out}\,\pi_1(S)$. In the case of the punctured torus, the form can be described explicitly: see [30] for further details. A bivector field dual to ω is given by

$$\frac{1}{2\pi}\left((2x - yz)\frac{\partial}{\partial y} \wedge \frac{\partial}{\partial z} + (2y - zx)\frac{\partial}{\partial z} \wedge \frac{\partial}{\partial x} + (2z - xy)\frac{\partial}{\partial x} \wedge \frac{\partial}{\partial y}\right).$$

The group $\mathrm{MCG}(S) \cong \mathrm{Out}\,\pi_1(S)$ is well known to be isomorphic to $GL_2\mathbb{Z}$. View the punctured torus S as a quotient of the Euclidean plane by two linearly independent translations, with the origin (and its orbit under the translations) removed. Then a homeomorphism of S is isotopic to a unique linear mapping which preserves the lattice, and we obtain a natural identification $\mathrm{Out}\,\pi_1(S) \cong \mathrm{MCG}(S) \cong GL_2\mathbb{Z}$.

Now $\mathrm{Aut}\,\pi_1(S)$ acts simply and transitively on bases, so $\mathrm{Out}\,\pi_1(S) \cong GL_2\mathbb{Z}$ acts simply and transitively on *conjugacy classes* of bases, i.e. bases subject to the equivalence relation $(G, H) \sim (AGA^{-1}, AHA^{-1})$ for $A \in \pi_1(S)$. Considering elements of $\pi_1(S)$ up to conjugacy has a geometric interpretation as curves in S up to to free homotopy. Simple closed curves up to free homotopy are determined by the pair (a, b), where $a \in \mathbb{Z}$ and $b \in \mathbb{Z}$ denote the number of times the curve traverses a chosen meridian and longitude of S. If we consider (a, b) as a vector, then an outer automorphism gives a matrix in $GL_2\mathbb{Z}$ which acts on (a, b) by the usual matrix multiplication. Considering a conjugacy class of *bases* also has a geometric interpretation, as a pair of basis curves up to a free homotopy of the pair of curves.

It's well known that $GL_2\mathbb{Z}$ has a small set of generators, for instance

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

This result has a nice interpretation geometrically. The modular group $PSL_2\mathbb{Z}$ is a discrete subgroup of $PSL_2\mathbb{R}$ and we may consider its action by isometries on the upper half plane model of $\mathbb{H}^2$. There is a natural fundamental domain for this action which tessellates the hyperbolic plane by hyperbolic ideal triangles. The two isometries

$$\pm \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}, \quad \pm \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

give edge pairings of the fundamental domain, hence generate $PSL_2\mathbb{Z}$. Taking two of their lifts into $SL_2\mathbb{Z}$, along with $-I$, we obtain a set of generators for $SL_2\mathbb{Z}$. And adding a single generator with determinant -1, we obtain a set of generators for $GL_2\mathbb{Z}$.

There is much more to say about these actions of $GL_2\mathbb{Z}$ and $PSL_2\mathbb{Z}$: they are useful in studying quadratic forms (e.g. [13], [48]), diophantine analysis (e.g. [45], [55]), complex dynamics (e.g. [6]) as well as hyperbolic surfaces (e.g. [7]).

It follows from the above that any two conjugacy classes of bases of $\pi_1(S)$ are related by some combination of the matrices above, considered as elements of $\text{Out}\, \pi_1(S) \cong \text{MCG}(S)$. We will consider the actions of these matrices on $X(S)$ separately.

(i) **The matrix** $M = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$**.**

As an element of the mapping class group, M is orientation-reversing and an involution. If S and G, H are in the standard position of figure 6.2, then M acts as a reflection in a plane of symmetry of S, preserving G but sending $H \mapsto H^{-1}$. In any case the automorphism $(G, H) \mapsto (G, H^{-1})$ descends to $M \in \text{Out}\, \pi_1(S)$. Letting G', H' be the image of G, H under the automorphism, and letting $(x, y, z) = (\text{Tr}\, g, \text{Tr}\, h, \text{Tr}\, gh)$, $(x', y', z') = (\text{Tr}\, g', \text{Tr}\, h', \text{Tr}\, g'h')$ denote the respective characters obtained, we have $(G', H', G'H') = (G, H^{-1}, GH^{-1})$ so it follows that

$$(x', y', z') = (\text{Tr}\, g', \text{Tr}\, h', \text{Tr}\, g'h') = (\text{Tr}\, g, \text{Tr}\, h^{-1}, \text{Tr}\, gh^{-1}) = (x, y, xy - z).$$

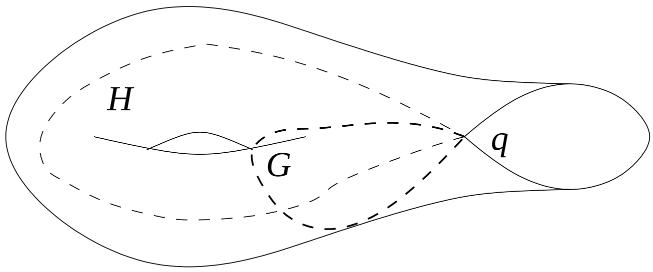

Figure 6.2: A standard set of basis curves on S

Here we use the trace relations given in section 4.4. Since $M = M^{-1}$, the actions of M and M^{-1} on $X(S)$ is given by

$$(x, y, z) \mapsto (x, y, xy - z).$$

(ii) **The matrix $-I$.**

As an element of the mapping class group, $-I$ gives a homeomorphism which is isotopic to an involution. If G and H are in the standard position of figure 6.2, then the homeomorphism acts as a rotation of π about the boundary, sending $G \mapsto G^{-1}$ and $H \mapsto H^{-1}$. In general the automorphism $(G, H) \mapsto (G^{-1}, H^{-1})$ is given in $\operatorname{Out} \pi_1(S)$ by $-I$. The induced action of $-I$ on the character variety $X(S)$ is trivial as $\operatorname{Tr} g = \operatorname{Tr} g^{-1}$, $\operatorname{Tr} h = \operatorname{Tr} h^{-1}$ and $\operatorname{Tr} g^{-1}h^{-1} = \operatorname{Tr} gh$.

(iii) **The matrix $M = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$.**

This matrix is of order 3. (Its image in $PSL_2\mathbb{R}$ considered as an isometry of the upper half plane is a rotation of $2\pi/3$ about $\frac{1}{2} + \frac{\sqrt{3}}{2}i$.) The automorphism $(G, H) \mapsto (H, H^{-1}G^{-1})$ descends to $M \in \operatorname{Out} \pi_1(S) \cong GL_2\mathbb{Z}$. Now if $(\operatorname{Tr} g, \operatorname{Tr} h, \operatorname{Tr} gh) = (x, y, z)$ then we have

$$(\operatorname{Tr} g', \operatorname{Tr} h', \operatorname{Tr} g'h') = (\operatorname{Tr} h, \operatorname{Tr} h^{-1}g^{-1}, \operatorname{Tr} g^{-1}) = (y, z, x).$$

So the actions of M, M^{-1} on $X(S)$ are given by

$$(x, y, z) \mapsto (y, z, x), (z, x, y).$$

(iv) **The matrix $M = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.**

The automorphism $(G, H) \mapsto (G, GH)$ descends to $M \in \operatorname{Out} \pi_1(S)$. As an element of the mapping class group M is a Dehn twist about G. We have $(x', y', z') = (\operatorname{Tr} g', \operatorname{Tr} h', \operatorname{Tr} g'h') = (\operatorname{Tr} g, \operatorname{Tr} gh, \operatorname{Tr} g^2h)$ which can easily be computed in terms of x, y, z. Similarly we have M^{-1} represented by the automorphism $(G, G^{-1}H)$ and can again compute $(x', y', z') = (\operatorname{Tr} g, \operatorname{Tr} g^{-1}h', \operatorname{Tr} h)$. The actions of M, M^{-1} on $X(S)$ are given by

$$(x, y, z) \mapsto (x, z, xz - y), (x, xy - z, y).$$

Thus we have obtained, in simple algebraic form, an equivalence relation on characters of representations $\rho : \pi_1(S) \longrightarrow SL_2\mathbb{R}$. If two representations ρ_1, ρ_2 are conjugate after a change of basis, then that change of basis is an automorphism which, up to conjugacy, can be considered as an element of $GL_2\mathbb{Z}$, hence can be written as a product of the generating matrices and the inverses considered above, hence the two characters are related by the equivalences described. Conversely, if two characters $(x, y, z), (x', y', z')$ are related by the equivalences described, then (provided $\kappa(x, y, z) = \kappa(x', y', z') \neq 2$ so ρ_1, ρ_2 are ireducible) the corresponding automorphisms give a change of basis for which ρ_1, ρ_2 have identical characters, hence by theorem 6.1.1 are conjugate representations.

The set of equivalences generating the equivalence relation on $X(S)$ is not in the most aesthetically pleasing form. To summarise we have the equivalence relation generated by

$$(x, y, z) \sim (x, y, xy - z),\ (y, z, x),\ (z, x, y),\ (x, z, xz - y),\ (x, xy - z, y).$$

Of these the third and fifth are clearly redundant as they are inverses of the second and fourth. Furthermore any permutation of the coordinates is permitted; for we have

$$\begin{aligned}(x, y, z) &\sim (x, xy - z, y) \\ &\sim (y, x, xy - z) \quad \text{by } (x, y, z) \sim (z, x, y) \\ &\sim (y, x, z) \quad \text{by } (x, y, z) \sim (x, y, xy - z).\end{aligned}$$

The two permutations $(x, y, z) \mapsto (y, z, x), (y, x, z)$ clearly generate the group of permutations on the coordinates. But this makes the fourth equivalence relation above redundant also. We can now record this result.

Proposition 6.3.1 *Let $\rho_1, \rho_2 : \pi_1(S) \longrightarrow SL_2\mathbb{R}$ be irreducible representations and let (G_1, H_1), (G_2, H_2) be free bases of $\pi_1(S)$ such that ρ_i has character (x_i, y_i, z_i) with respect to the basis (G_i, H_i). The following are equivalent:*

(i) There exists $\phi \in \operatorname{Aut} \pi_1(S)$ such that $\rho_1 \circ \phi$ and ρ_2 are conjugate representations into $SL_2\mathbb{R}$;

(ii) $(x, y, z) \sim (x', y', z')$ *under the equivalence relation generated by permutation of coordinates and the relation* $(x, y, z) \sim (x, y, xy - z)$. ∎

For representations into $PSL_2\mathbb{R}$ we must also add the sign-change relations described in section 6.1. We obtain the following proposition.

Proposition 6.3.2 *Let* $\rho_1, \rho_2 : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ *be irreducible representations and let* (G_1, H_1), (G_2, H_2) *be free bases of* $\pi_1(S)$. *Choosing lifts of* g_i, h_i *into* $SL_2\mathbb{R}$ *arbitrarily, let* ρ_i *have character* (x_i, y_i, z_i) *with respect to the basis* (G_i, H_i). *The following are equivalent:*

(i) There exists $\phi \in \mathrm{Aut}\, \pi_1(S)$ *such that* $\rho_1 \circ \phi$ *and* ρ_2 *are conjugate representations in* $PSL_2\mathbb{R}$*;*

(ii) $(x, y, z) \sim (x', y', z')$ *under the equivalence relation generated by permutation of coordinates and the relations* $(x, y, z) \sim (x, y, xy - z)$ *and* $(x, y, z) \sim (-x, -y, z)$. ∎

Triples of numbers with this relation are known as *Markoff triples*. Note the equivalence relation can be considered here as the orbit space under the action of a certain group which we denote Γ, following [30], where

$$\Gamma \cong PGL_2\mathbb{Z} \ltimes \left(\frac{\mathbb{Z}}{2} \oplus \frac{\mathbb{Z}}{2} \right).$$

This is a semidirect prouct. The $PGL_2\mathbb{Z}$ arises since the action of $-I$ is trivial; the $\mathbb{Z}/2 \oplus \mathbb{Z}/2$ corresponds to the sign-change relations.

If we restrict our attention to orientation-preserving changes of basis, then we may not consider all of the moves above. In particular, we cannot transpose coordinates freely. But we can certainly apply the relations given by the action of matrices (ii)–(iv) above; and if we are considering representations into $PSL_2\mathbb{R}$, then we may apply the sign-change relations also. While at times we will need to consider the orientation of a basis, we will always consider the above machinery without orientation-preserving restrictions.

6.4 The virtually abelian representations

Another class of degenerate representations is the class of virtually abelian representations. Recall that a representation is *abelian* if it has abelian image. A representation is *virtually abelian* if its image contains an abelian subgroup of finite index.

First we investigate abelian representations. We have seen above that all abelian representations are reducible. Conversely, a character of a reducible representation is also the character of an abelian representation. For a reducible representation ρ can be taken to map G, H to upper triangular matrices. We can then define a representation ρ' taking G, H to diagonal matrices, simply ignoring the top right entry. This defines an abelian representation with the same character.

It's easy to see that the image of an abelian representation $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ consists of one of the following:

(i) elliptics which all rotate about the same point, and the identity;

(ii) parabolics with the same fixed point at infinity, and the identity;

(iii) hyperbolics with the same axis, and the identity;

(iv) the identity alone.

We will now give a characterisation of virtually abelian representations. First, define a set V as

$$V = \{0 \times 0 \times \mathbb{R}\backslash[-2,2]\} \cup \{0 \times \mathbb{R}\backslash[-2,2] \times 0\} \cup \{\mathbb{R}\backslash[-2,2] \times 0 \times 0\}.$$

We can easily verify, checking the conditions of theorem 6.1.2, that $V \subset X(S)$. Further, using 6.1.2 or figure 6.1, the set of points in $X(S)$ with two coordinates equal to 0 is precisely V, taken together with the six points $(0,0,\pm 2)$, $(0,\pm 2,0)$, $(\pm 2,0,0)$. (If $(0,0,z) \in X(S)$ then $\kappa(x,y,z) = z^2 - 2 \geq 2$ is equivalent to $|z| \geq 2$.) We can also see from above that no abelian representations have characters in V.

A geometric description of representations with characters in V can easily be given.

Lemma 6.4.1 *Let $g, h \in PSL_2\mathbb{R}$. The following are equivalent:*

(i) We may lift g, h to $SL_2\mathbb{R}$ so that $(\mathrm{Tr}(g), \mathrm{Tr}(h), \mathrm{Tr}(gh)) \in V$.

(ii) Two of $\{g, h, gh\}$ are half-turns about points $q_1 \neq q_2 \in \mathbb{H}^2$ and the third is a nonzero translation along the axis q_1q_2.

Proof Half-turns have trace 0 and nonzero translations (i.e. hyperbolic isometries) have trace greater than 2 in magnitude, so clearly if $\{g, h, gh\}$ are isometries of the required type then $(\mathrm{Tr}(g), \mathrm{Tr}(h), \mathrm{Tr}(gh)) \in V$.

Now suppose $\mathrm{Tr}(g) = \mathrm{Tr}(h) = 0$ and $|\mathrm{Tr}(gh)| > 2$. We will only deal with this case; the other cases are similar. Then g, h are half-turns about two points q_1, q_2. If

$q_1 = q_2$ then $gh = 1$ and $\text{Tr}(gh) = \pm 2$, a contradiction. So $q_1 \neq q_2$, and both g, h preserve the line $q_1 q_2$, reversing its orientation. The composition gh therefore also preserves the line $q_1 q_2$ but maintains its orientation. Since $q_1 \neq q_2$ we have $gh \neq 1$, and therefore gh must be a nonzero translation along $q_1 q_2$. ∎

Note that in this situation, the subgroup $\langle g, h \rangle$ of $PSL_2\mathbb{R}$ is an infinite dihedral group consisting of translations along $q_1 q_2$ and half-turns about points on $q_1 q_2$. It therefore contains an index 2 subgroup of translations along $q_1 q_2$, which is abelian. So ρ in this case is indeed virtually abelian.

Lemma 6.4.2 *Let ρ be a representation with $(\text{Tr}(g), \text{Tr}(h), \text{Tr}(gh)) \in V$. Let G', H' be another basis of $\pi_1(S)$. Then $(\text{Tr}(g'), \text{Tr}(h'), \text{Tr}(g'h')) \in V$ also.*

Proof We have $(\text{Tr}\, g, \text{Tr}\, h, \text{Tr}\, gh) \in V$. Now recall that the action induced by the mapping class group $GL_2\mathbb{Z}$ on the space $X(S)$ of characters of representations is generated by the transformations $(x, y, z) \mapsto (x, y, xy - z)$, the permutations or coordinates, and sign changes. It's easy to check that these transformations send points of V to points of V. ∎

We now show that this is a complete characterisation of virtually abelian groups.

Lemma 6.4.3 *Let $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ be a representation. The character*

$$(x, y, z) = (\text{Tr}\, g, \text{Tr}\, h, \text{Tr}\, gh) \in V$$

if and only if ρ is virtually abelian but not abelian.

Proof We have already established that $(x, y, z) \in V$ is the character of a virtually abelian but not abelian representation. So let $\Lambda = \rho(\pi_1(S)) \subset PSL_2\mathbb{R}$ be virtually abelian but not abelian. So there is a finite index subgroup F of Λ which is abelian. Let F have index $n > 1$ in Λ. Note if $\alpha, \beta \in \Lambda$ lie in the same left coset of F then $\alpha F = \beta F$ and $F\alpha^{-1} = F\beta^{-1}$, so $\alpha F \alpha^{-1} = \beta F \beta^{-1}$. Hence there are only finitely many conjugate subgroups of F in Λ; by taking their intersection we obtain a *normal* finite index abelian subgroup of Λ. Passing to this subgroup, we may assume F is normal.

Let $\text{Fix}(F)$ denote the set of points in $\overline{\mathbb{H}^2}$ fixed by every element of F. I claim $\text{Fix}(F)$ is invariant under the action of Λ. Take $h \in \Lambda$ and $p \in \text{Fix}(F)$; we must show $h(p) \in \text{Fix}(F)$. So take $f \in F$; then $h^{-1} f h \in F$ by normality; so $h^{-1} f h(p) = p$, hence $f(h(p)) = h(p)$. So $h(p) \in \text{Fix}(F)$ as desired. We now split into cases according to the possibilities for F.

Case (i). Suppose $F = \{1\}$, so Λ is finite, so every element has finite order, hence is elliptic or the identity. Take an arbitrary point $q \in \mathbb{H}^2$ and let p be the centre of mass of the (finite) orbit of q under Λ (see [60, 2.5.19] for more details). Then p is fixed by every element of Λ. So every element of Λ is the identity or an elliptic fixing p. Hence Λ is abelian, a contradiction.
Case (ii). Assume F consists of the identity and elliptics fixing a point q, so $\text{Fix}(F) = q$. So every element of Λ fixes q, and Λ consists of the identity and elliptics fixing q. Thus Λ is abelian, a contradiction.
Case (iii). Assume F consists of the identity and parabolics with fixed point q, so $\text{Fix}(F) = q$. So every element of Λ fixes q. There cannot exist a hyperbolic $h \in \Lambda$, for then $h^n \in F$ would be hyperbolic. So Λ consists of the identity and parabolics fixing q, a contradiction.
Case (iv). Now assume F consists of the identity and hyperbolic isometries with axis l, so $\text{Fix}(F) = \bar{l}$ (consisting of l and its endpoints at infinity). So every element of Λ fixes $\bar{l}$, hence is either the identity, or hyperbolic with axis l, or elliptic of order 2 with fixed point on l. If there are no elliptics then Λ is abelian and we have a contradiction. Otherwise the the translations (and the identity) form an index-2 subgroup of Λ. The pair g, h (where $G, H \in \pi_1(S)$ is a basis) must contain at least one half turn; hence the triple g, h, gh contains exactly two half turns about distinct points on l, and one hyperbolic element translating along l. By lemma 6.4.1 the character of ρ with respect to this basis lies in V. ■

6.5 The reducible representations

We have seen (proposition 6.1.3) that the reducible representations are precisely those with $\text{Tr}[g, h] = 2$. We can classify these more explicitly. These include abelian representations. From the previous section, all the representations which are virtually abelian, but not abelian, have character in V, hence have $\text{Tr}[g, h] > 2$. So we have immediately:

Lemma 6.5.1 *A reducible virtually abelian representation $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ is abelian.* ■

We will now describe the non-abelian reducible representations rather explicitly.

Lemma 6.5.2 *Let ρ be a non-abelian reducible representation $\pi_1(S) \longrightarrow PSL_2\mathbb{R}$ and let G, H be a basis of $\pi_1(S)$. Then one of the following occurs:*

(i) one of g, h is hyperbolic and the other is parabolic, and g, h share a fixed point at infinity;

(ii) g, h are both hyperbolic, sharing exactly one fixed point at infinity.

PROOF If g or h is the identity then ρ is trivially abelian. Suppose g is elliptic. Then we may conjugate in $PSL_2\mathbb{R}$ so that the fixed point of g lies at i in the upper half plane model. Then we may write

$$g = \pm \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad h = \pm \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

where $\sin\theta \neq 0$. We obtain

$$\text{Tr}[g, h] = 2 + (a^2 + b^2 + c^2 + d^2 - 2)\sin^2\theta$$

hence as $\text{Tr}[g, h] = 2$ and $\sin\theta \neq 0$,

$$a^2 + b^2 + c^2 + d^2 = 2 = 2(ad - bc).$$

Thus $(a - d)^2 + (b + c)^2 = 0$, so $a = d$ and $b = -c$. So h takes the form

$$h = \pm \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

and $a^2 + b^2 = 1$, so for some ϕ we have

$$h = \pm \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}.$$

That is, h is also a rotation about i, so g, h commute and ρ is abelian.

If h is elliptic, then we apply the same argument noting $\text{Tr}[h, g] = \text{Tr}[g, h]$.

Hence each of g, h is hyperbolic or parabolic. Suppose first that one of g, h is parabolic, without loss of generality g. Then we may conjugate in $PSL_2\mathbb{R}$, and replacing G with G^{-1} if necessary we have

$$g^{\pm 1} = \pm \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad h = \pm \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

We can calculate $\text{Tr}[g, h] = \text{Tr}[g^{-1}, h] = 2 + c^2$ but ρ is reducible, so $\text{Tr}[g, h] = 2$. Thus $c = 0$ and h is upper triangular, hence h fixes ∞ in common with g. If h is parabolic then ρ is abelian, since g, h are parabolics with the same fixed point. So h is hyperbolic.

Suppose now that both g, h are hyperbolic. Note g, h cannot share two fixed points at infinity, as ρ is not abelian. Hence we may conjugate so that g has fixed points at infinity $\{-1, 1\}$ in the upper half-plane model, and h has fixed points $\{r, \infty\}$. Then for some real θ, ϕ we have

$$g = \pm \begin{bmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{bmatrix}, \quad h = \pm \begin{bmatrix} e^{\phi} & -2r\sinh\phi \\ 0 & e^{-\phi} \end{bmatrix}.$$

We calculate

$$2 = \mathrm{Tr}[g, h] = 2 + 4(r^2 - 1)\sinh^2\theta\ \sinh^2\phi$$

so $(r^2 - 1)\sinh^2\theta\ \sinh^2\phi = 0$. If $\sinh\theta = 0$ or $\sinh\phi = 0$ then respectively g or h is the identity and ρ is abelian; therefore $r = \pm 1$, and g, h share exactly one fixed point at infinity. ■

Chapter 7

The Construction of Punctured Tori

7.1 Statement and preliminaries

Throughout this chapter, as usual, let S be a punctured torus, and let G, H be a basis of $\pi_1(S)$, with a basepoint q chosen on the boundary. Let $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ be a representation, and let $\rho(G) = g, \rho(H) = h$. We prove the following result.

Theorem A *A representation $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ is the holonomy of a hyperbolic cone-manifold structure on S with at most one corner point and no interior cone points if and only if ρ is not virtually abelian.*

As discussed previously, we may take some lift of g, h, gh into $SL_2\mathbb{R}$ and let $(x, y, z) = (\operatorname{Tr} g, \operatorname{Tr} h, \operatorname{Tr} gh) \in X(S)$ be the character of ρ, which is well-defined up to the equivalence relations $(x, y, z) \sim (-x, -y, z) \sim (-x, y, -z) \sim (x, -y, -z)$. Then we have $\operatorname{Tr}[g, h] = \kappa(x, y, z) = x^2 + y^2 + z^2 - xyz - 2$, which is well-defined regardless of the choice of lift into $SL_2\mathbb{R}$; indeed $[g, h]$ gives a well-defined element of $\widetilde{PSL_2}\mathbb{R}$. The proof is split into cases according to the value of $\operatorname{Tr}[g, h]$.

In section 7.2 we treat the case $\operatorname{Tr}[g, h] \in (-\infty, -2)$. By corollary 3.7.3, we see $\operatorname{Tr}[g, h] < -2$ implies that $[g, h] \in \operatorname{Hyp}_1 \cup \operatorname{Hyp}_{-1}$. By proposition 4.3.4, we have the relative Euler class well-defined and $\mathcal{E}(\rho)[S] = \pm 1$. We will construct a hyperbolic cone-manifold structure with a preferred orientation, accordingly as $[g, h] \in \operatorname{Hyp}_1$ or Hyp_{-1}.

In section 7.3 we treat $\operatorname{Tr}[g, h] = -2$. In this case we have, similarly, from corollary 3.7.3, $[g, h] \in \operatorname{Par}_1^-$ or $\operatorname{Par}_{-1}^+$, and from proposition 4.3.4, $\mathcal{E}(\rho)[S] = \pm 1$. Again we will construct a hyperbolic cone-manifold structure with a preferred orientation accordingly as $[g, h] \in \operatorname{Par}_1^-$ or $\operatorname{Par}_{-1}^+$.

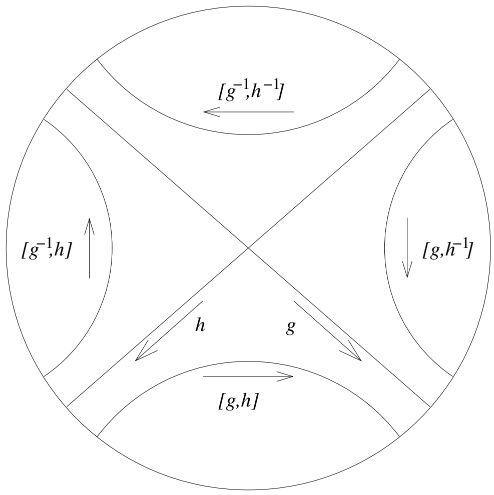

Figure 7.1: The arrangement of axes of commutators if $\mathrm{Tr}[g,h] < -2$,

In section 7.4 we consider $\mathrm{Tr}[g,h] \in (-2,2)$. In this case from corollary 3.7.3 we have $[g,h] \in \mathrm{Ell}_{-1}$ or Ell_1, but the relative Euler class is not well-defined. We will find cone-manifold structures of one of the two possible orientations accordingly as $[g,h] \in \mathrm{Ell}_1$ or Ell_{-1}; and there are "large angle" or "small angle" cases as described in chapter 5.

In section 7.5 we treat the case $\mathrm{Tr}[g,h] = 2$. From corollary 3.7.3 we have $[g,h] \in \{1\} \cup \mathrm{Par}_0$. By proposition 6.1.3, these are precisely the reducible representations, and by proposition 4.3.4 we have $\mathcal{E}(\rho)[S] = 0$. Some of these representations are virtually abelian (in fact abelian, using lemma 6.4.3); we will prove these are not holonomy representations. For the other reducible representations we will find a cone-manifold structure of a preferred orientation, accordingly as $[g,h] \in \mathrm{Par}_0^+$ or Par_0^-.

In section 7.6 we consider the most difficult case, $\mathrm{Tr}[g,h] > 2$. Some of these representations are virtually abelian, and we will eliminate these. For the other representations, by corollary 3.7.3 we have $[g,h] \in \mathrm{Hyp}_0$, and by proposition 4.3.4 we have $\mathcal{E}(\rho)[S] = 0$. There is no preferred orientation; we do not specify it in advance.

7.2 The case $\mathrm{Tr}[g,h] < -2$: complete and discrete

From lemma 3.2.2, if $\mathrm{Tr}[g,h] < 2$ then g, h are both hyperbolic and their axes cross. If $\mathrm{Tr}[g,h] < -2$ then this commutator is hyperbolic. By four applications of lemma 3.2.4, the arrangement of axes of various commutators is as shown in figure 7.1.

Taking an arbitrary point p on the axis of $[g,h]$ we investigate the arrange-

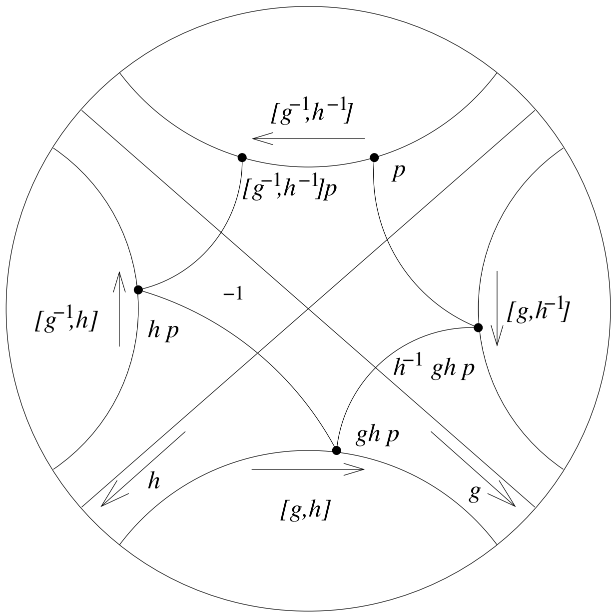

Figure 7.2: $\mathcal{P}(g,h,p)$ bounds an embedded disc.

ment of $\mathcal{P}(g,h,p)$. In general we have $\alpha(\mathrm{Axis}\,\beta) = \mathrm{Axis}(\alpha\beta\alpha^{-1})$. So hp lies on $\mathrm{Axis}[h,g^{-1}] = \mathrm{Axis}[g^{-1},h]$. Similarly, $ghp \in \mathrm{Axis}[g,h]$ and $h^{-1}gh \in \mathrm{Axis}[g,h^{-1}]$. Obviously $[g^{-1},h^{-1}]p \in \mathrm{Axis}[g^{-1},h^{-1}]$, as a translation of p. Given the arrangement of axes, it's clear that $\mathcal{P}(g,h,p)$ bounds an embedded disc. See figure 7.2.

By lemma 5.1.5, this gives a hyperbolic cone-manifold structure on S with no interior cone points and one corner point. Now $[g^{-1},h^{-1}]$, as a hyperbolic isometry, simply translates along $\mathrm{Axis}[g^{-1},h^{-1}]$. So $\mathrm{Tw}([g^{-1},h^{-1}],p)$ is a multiple of 2π. Since $[g,h] \in \mathrm{Hyp}_{\pm 1}$ and $[g,h]$ is conjugate to $[g^{-1},h^{-1}]$, we have $\mathrm{Tw}([g^{-1},h^{-1}],p) = \pm 2\pi$. By lemma 5.4.1, we have a cone angle $\theta \equiv \pi \bmod 2\pi$. By Gauss-Bonnet, proposition 2.2.2, we have $\theta \in (0, 3\pi)$. So $\theta = \pi$. That is, the corner point at q is actually no corner at all, and we have obtained a hyperbolic structure on S with totally geodesic boundary.

Note the many choices here: we can take any basis g,h for $\pi_1(S)$, and any point p on $\mathrm{Axis}[g^{-1},h^{-1}]$. It is clear why: as the holonomy of a complete hyperbolic structure on S, ρ is discrete, and S is just the quotient of the convex core of ρ by the action of the holonomy group. This underlying surface is independent of our choice of basis, and we are free to choose any basepoint on $\mathrm{Axis}[g^{-1},h^{-1}]$. The quotient of $\mathbb{H}^2$ by the holonomy group is the underlying "flared" surface discussed in chapter 5 and depicted in figure 5.13. By choosing p inside or outside the convex core, we may extend or truncate the surface with geodesic boundary, as described in chapter 5. The axes of g and h, for p sufficiently close to $\mathrm{Axis}[g^{-1},h^{-1}]$, intersect in the interior of $\mathcal{P}(g,h,p)$ and project to the simple closed geodesics in S which are

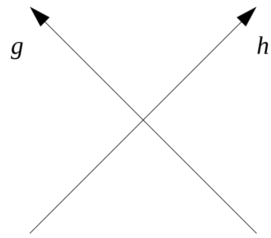

Figure 7.3: Axes of g, h intersecting

the shortest in the free homotopy class of G, H.

Proposition 7.2.1 *Let ρ be a representation and G, H a basis of $\pi_1(S)$ with $\mathrm{Tr}[g,h] < -2$. Suppose $[g,h] \in \mathrm{Hyp}_1$ (resp. Hyp_{-1}). Then:*

(i) ρ is the holonomy of a complete hyperbolic structure in which ∂S, traversed in the direction homotopic to $[G, H]$, bounds S on its left (resp. right).

(ii) The axes of g, h intersect in the manner shown in figure 7.3 (resp. the opposite manner).

(iii) For $p \in \mathrm{Axis}[g^{-1}, h^{-1}]$ we have $\mathrm{Tw}([g^{-1}, h^{-1}], p) = 2\pi$ (resp. -2π).

(iv) For p sufficiently close to $\mathrm{Axis}[g^{-1}, h^{-1}]$, we obtain a hyperbolic cone-manifold structure on S with one corner point. The corner angle is given by $\theta = 3\pi - \mathrm{Tw}([g^{-1}, h^{-1}], p)$ (resp. $3\pi + \mathrm{Tw}([g^{-1}, h^{-1}], p)$).

Proof Suppose the axes of g, h intersect in the manner of figure 7.3. Taking a point $p \in \mathrm{Axis}[g^{-1}, h^{-1}]$, we obtain the situations of figures 7.1 and 7.2. We see that the segment $p \to [g^{-1}, h^{-1}]p$ bounds $\mathcal{P}(g,h,p)$ to its left; hence ∂S traversed in the direction of $[G, H]$ bounds S on its left. Chasing unit vectors in the manner of the proof of lemma 5.4.1, since we have an explicit arrangement of the axes of g, h, we may observe the twist of $[g^{-1}, h^{-1}]$ at p. Consider a lift $\tilde{g}$ of g to $\widetilde{PSL_2\mathbb{R}}$ which simply flows unit vectors along the constant distance curves from $\mathrm{Axis}\, g$, and similarly for h. We see that in the situation of 7.1, starting from p, a unit vector is flowed anticlockwise to hp by $\tilde{h}$; then anticlockwise again to ghp by $\tilde{g}$; and again to $h^{-1}ghp$ and $[g^{-1}, h^{-1}]p$. So the twist of $[g^{-1}, h^{-1}]$ at p is anticlockwise, i.e. positive. Above we showed $\mathrm{Tw}([g^{-1}, h^{-1}], p) = \pm 2\pi$, so we conclude $\mathrm{Tw}([g^{-1}, h^{-1}], p) = 2\pi$. We knew that $[g, h] \in \mathrm{Hyp}_1$ or Hyp_{-1}; since the twist at p of its conjugate $[g^{-1}, h^{-1}]$ is positive, from proposition 3.4.4 we have $[g, h] \in \mathrm{Hyp}_1$.

If the axes of g, h intersect in the opposite manner, then we similarly conclude all the respective statements.

By continuity, as discussed at length in chapter 5, choosing p sufficiently close to $\mathrm{Axis}[g^{-1}, h^{-1}]$ we obtain the desired cone-manifold structure. The corner angle θ varies continuously with $\mathrm{Tw}([g^{-1}, h^{-1}], p)$, obeying the relation of lemma 5.4.1. If the segment $p \to [g^{-1}, h^{-1}]p$ bounds $\mathcal{P}(g,h,p)$ on its left, i.e. the convex core lies to the left of this directed segment, then perturbing p outside the convex core increases $\mathrm{Tw}([g^{-1}, h^{-1}], p)$, and hence decreases θ. On the other hand, perturbing p inside the convex core increases θ. Since we have $[g,h] \in \mathrm{Hyp}_1$, $\mathrm{Tw}([g^{-1}, h^{-1}], p) \in (\pi, 3\pi)$. Since (by lemma 5.4.1) $\theta \equiv \pi - \mathrm{Tw}([g^{-1}, h^{-1}], p) \bmod 2\pi$, with $\theta = \pi$ when $\mathrm{Tw}([g^{-1}, h^{-1}], p) = 2\pi$, by continuity we must have $\theta = 3\pi - \mathrm{Tw}([g^{-1}, h^{-1}], p)$. Alternatively if $p \to [g^{-1}, h^{-1}]p$ bounds $\mathcal{P}(g,h,p)$ on its right, then perturbing p outside the convex core decreases $\mathrm{Tw}([g^{-1}, h^{-1}], p)$ and decreases θ; and conversely. We have $\theta \in (0, 2\pi)$, $\mathrm{Tw}([g^{-1}, h^{-1}], p) \in (-3\pi, -\pi)$, so $\theta = 3\pi + \mathrm{Tw}([g^{-1}, h^{-1}], p)$ in this case. ∎

Thus the description of chapter 5 holds true.

7.3 The case $\mathrm{Tr}[g,h] = -2$: parabolics and cusps

This case proceeds similarly to the previous case. By corollary 3.7.3, $[g,h]$ lies in Par_1^- or Par_{-1}^+. The isometries g, h are still hyperbolic and their axes cross. Using lemma 3.2.5 four times, we obtain the situation of figure 7.4.

Let r be the fixed point at infinity of $[g^{-1}, h^{-1}]$. Then, noting that $\alpha(\mathrm{Fix}\,\beta)$ is just $\mathrm{Fix}(\alpha\beta\alpha^{-1})$, we see that $hr = \mathrm{Fix}[g^{-1}, h]$, $ghr = \mathrm{Fix}[g,h]$, and $h^{-1}ghr = \mathrm{Fix}[g, h^{-1}]$. Obviously $[g^{-1}, h^{-1}]r = r$. So if we choose $p = r$ then $\mathcal{P}(g,h,p)$ is an ideal quadrilateral, with degenerate "boundary" edge, and bounds an embedded disc, as in figure 7.5.

We have obtained a hyperbolic structure on S, where the boundary has become a cusp. It is a complete hyperbolic structure without any cone points, and ρ is a discrete representation. Thus, as in the previous case, the quotient of $\mathbb{H}^2$ by the holonomy group forms an "underlying" cusped surface. If we take $p \in \mathbb{H}^2$, rather than at infinity, then we truncate this underlying surface and obtain a cone-manifold structure on S with no interior cone points and one corner point. Thus, as in the previous case, we can take any basis G, H and obtain an isometric hyperbolic surface.

In figure 7.4 we have drawn horocycles along which $[g^{-1}, h^{-1}]$ and its conjugates translate. From the discussion of lemma 3.4.3, as p approaches the fixed point of $[g^{-1}, h^{-1}]$, since $[g^{-1}, h^{-1}] \in \mathrm{Par}_1^- \cup \mathrm{Par}_{-1}^+$, $\mathrm{Tw}([g^{-1}, h^{-1}], p) \longrightarrow \pm 2\pi$. Thus by lemma 5.4.1, the corner angle θ is close to π or 3π. But the area of the pentagon

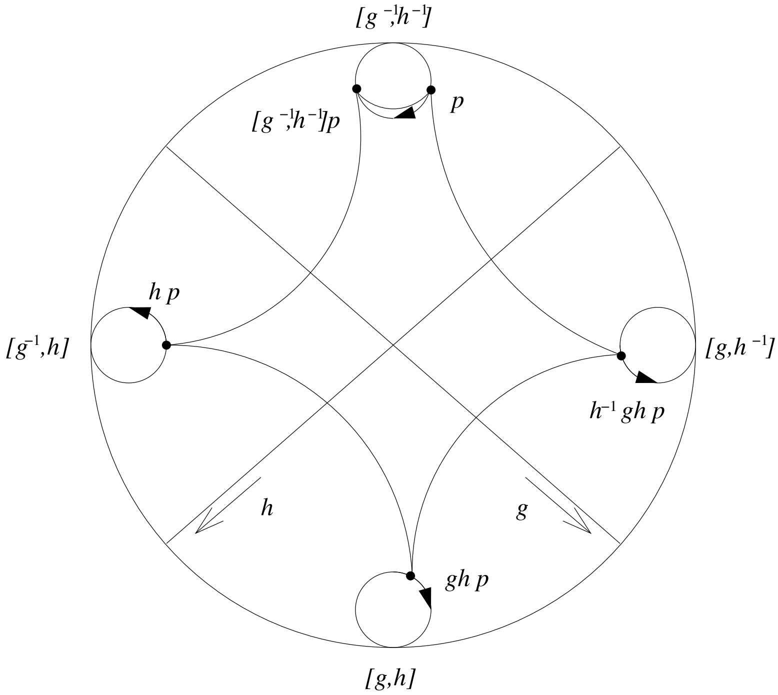

Figure 7.4: The situation when $\mathrm{Tr}[g,h] = -2$.

is $3\pi - \theta$, and the area is not close to 0; thus as p approaches the fixed point of $[g^{-1}, h^{-1}]$, $\theta \longrightarrow \pi$.

Proposition 7.3.1 *Let ρ be a representation and G, H a basis of $\pi_1(S)$ with $\mathrm{Tr}[g,h] = -2$. Suppose $[g,h] \in \mathrm{Par}_1^-$ (resp. Par_{-1}^+). Then:*

(i) ρ is the holonomy of a hyperbolic cone-manifold structure on S with no interior cone points and one corner point and ∂S, traversed in the direction homotopic to $[G,H]$, bounds S on its left (resp. right);

(ii) The axes of g, h intersect in the manner shown in figure 7.3 (resp. the opposite manner).

(iii) as p approaches the fixed point at infinity of $[g^{-1}, h^{-1}]$, $\mathrm{Tw}([g^{-1}, h^{-1}], p)$ approaches 2π from below (resp. -2π from above);

(iv) The corner angle is given by $\theta = 3\pi - \mathrm{Tw}([g,h], p)$ (resp. $3\pi + \mathrm{Tw}([g,h], p)$).

Proof Essentially identical to the proof in the hyperbolic case. By chasing unit vectors we see that in the case shown, $\mathrm{Tw}([g^{-1}, h^{-1}], p) > 0$, hence by lemma 3.4.4 $[g^{-1}, h^{-1}], [g,h] \in \mathrm{Par}_1^-$, so $\mathrm{Tw}([g^{-1}, h^{-1}], p) \in (\pi, 2\pi)$. So $[g^{-1}, h^{-1}]$ translates clockwise around the circle at infinity, and hence the geodesic segment $p \to [g^{-1}, h^{-1}]p$ bounds the pentagon on its left; hence ∂S traversed according to $[G,H]$ bounds S

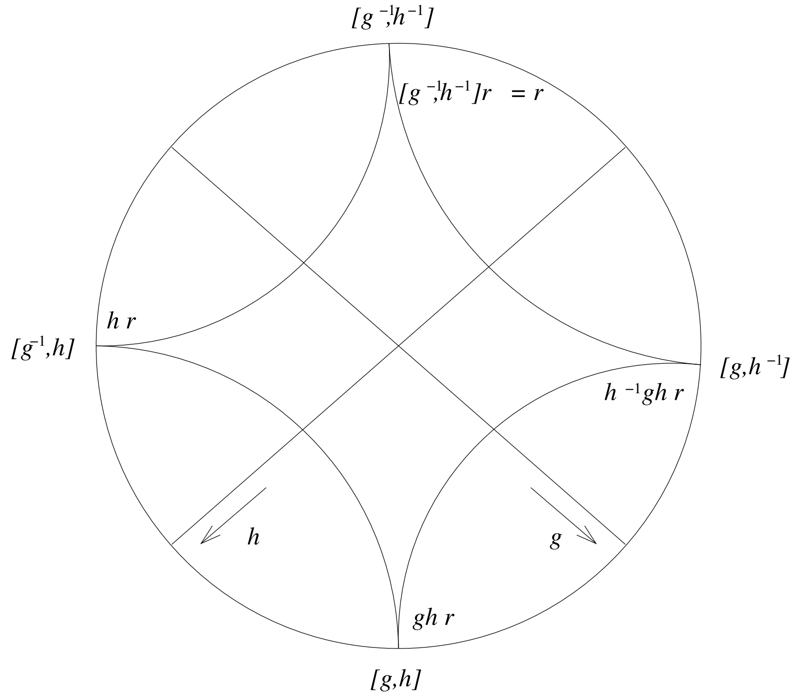

Figure 7.5: An ideal quadrilateral obtained by taking $p = \mathrm{Fix}[g^{-1}, h^{-1}]$.

on its left. For p close to the fixed point we then have $\mathrm{Tw}([g^{-1}, h^{-1}], p)$ close to 0 (mod 2π), by the discussion of lemma 3.4.3, and in $(\pi, 2\pi)$. Thus as p approaches the fixed point, $\mathrm{Tw}([g^{-1}, h^{-1}], p)$ approaches 2π from below.

If the axes of g, h intersect in the opposite manner, again we conclude all the respective statements.

The corner angle θ is given by the sum of the angles in the pentagon, and lemma 5.4.1 relates θ to $\mathrm{Tw}([g^{-1}, h^{-1}], p)$. Suppose first that $[g, h] \in \mathrm{Par}_1^-$, so that for p near the fixed point of $[g^{-1}, h^{-1}]$ we obtain $\mathcal{P}(g, h, p)$ with the directed segment $p \to [g^{-1}, h^{-1}]p$ bounding the pentagon on its left. As p approaches the fixed point of $[g^{-1}, h^{-1}]$, $\mathrm{Tw}([g^{-1}, h^{-1}], p)$ approaches 2π from below. Thus $\pi - \mathrm{Tw}([g^{-1}, h^{-1}], p)$ approaches $-\pi$ from above, hence θ approaches π from above, modulo 2π. But since by Gauss-Bonnet the angle in a pentagon lies in $(0, 3\pi)$, the angle approaches π (exactly) from above. Hence $\theta = 3\pi - \mathrm{Tw}([g^{-1}, h^{-1}], p)$, and by continuity this is true for all p near $\mathrm{Fix}[g^{-1}, h^{-1}]$. If the other situation arises, i.e. $[g, h] \in \mathrm{Par}_{-1}^+$, then as p approaches the fixed point of $[g^{-1}, h^{-1}]$, $\mathrm{Tw}([g^{-1}, h^{-1}], p)$ approaches -2π from above, so $\pi + \mathrm{Tw}([g^{-1}, h^{-1}], p)$ approaches $-\pi$ from above, so by a similar argument θ approaches π from above, given by $\theta = 3\pi + \mathrm{Tw}([g^{-1}, h^{-1}], p)$. ∎

That is, the further out to infinity we choose p, the "flatter" the corner angle obtained. This accords with our geometric intuition, as in figure 7.6.

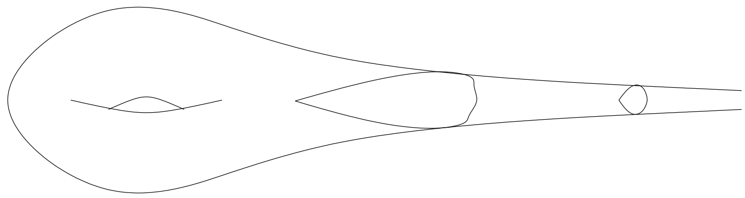

Figure 7.6: The corner angle $\theta \in (\pi, 2\pi)$, and tends to π as we truncate closer to infinity.

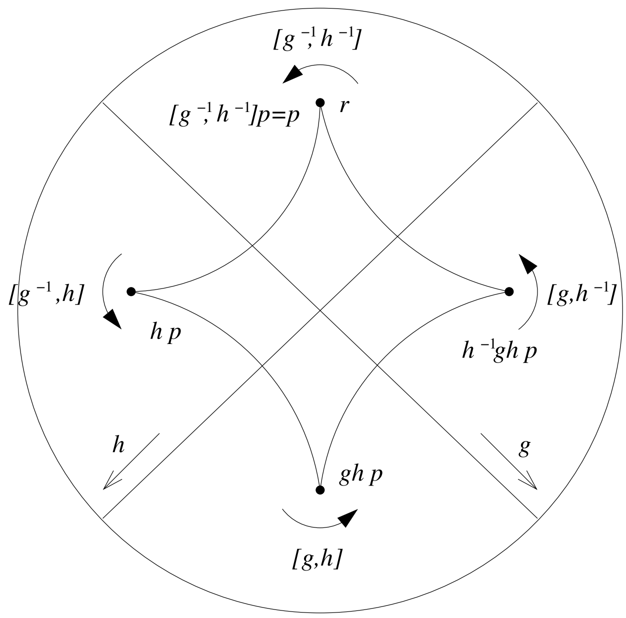

Figure 7.7: The situation when $\mathrm{Tr}[g,h] \in (-2,2)$ and $p = r$.

7.4 The case $\mathrm{Tr}[g,h] \in (-2,2)$

We commence our analysis of the case $-2 < \mathrm{Tr}[g,h] < 2$ similarly to the previous cases. By lemma 3.2.6 the fixed points of $[g,h]$, $[g^{-1},h]$, $[g,h^{-1}]$ and $[g^{-1},h^{-1}]$ are as shown in figure 7.7. Let r denote the fixed point of $[g^{-1},h^{-1}]$.

Letting $p = r$, the pentagon $\mathcal{P}(g,h,p)$ degenerates with the boundary shrinking to a point, so we obtain a quadrilateral with edges identified. Thus ρ is the holonomy of a hyperbolic cone-manifold structure on a punctured torus, where the boundary has been pinched to a point, as discussed in chapter 5.

We would like to perturb p away from the fixed point r, in a direction so that $\mathcal{P}(g,h,p)$ bounds an embedded (or immersed) disc, as illustrated in chapter 5. Recall the picture varied in a "large angle" or "small angle" case. For sufficiently small ϵ, we will perturb p around a small circle $C_\epsilon(r)$ of radius ϵ about r. The question is: how far can we perturb p around $C_\epsilon(r)$ so that $\mathcal{P}(g,h,p)$ is a non-degenerate pentagon bounding an embedded disc?

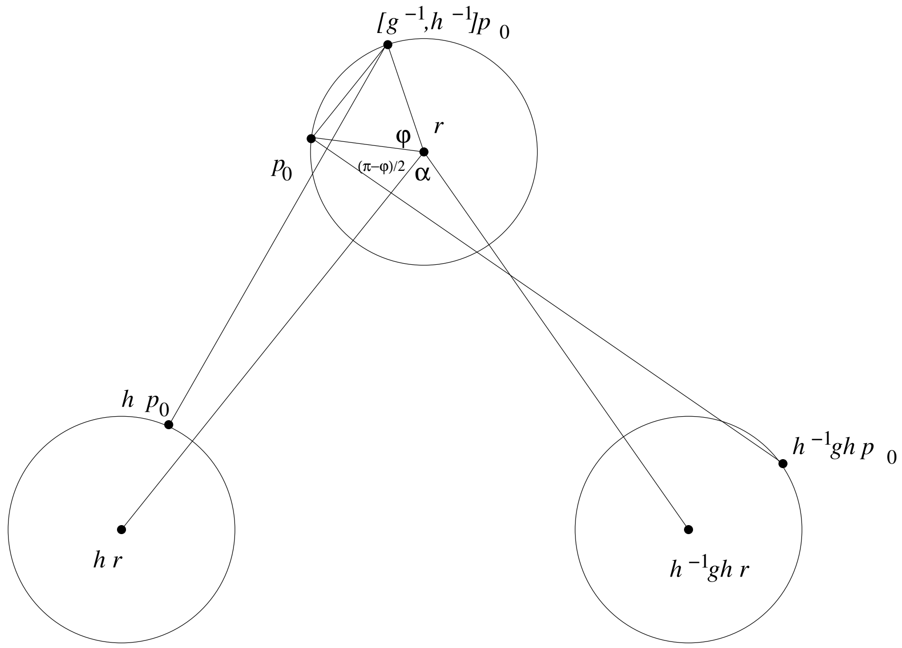

Figure 7.8: The situation in $\mathcal{P}(g,h,p)$ for $p \in C_\epsilon(r)$. (These lines are all hyperbolic geodesics.)

First a remark about orientation. From corollary 3.7.3 we know $[g,h] \in \mathrm{Ell}_{-1} \cup \mathrm{Ell}_1$. In the situation of figure 7.7, a simple unit vector chase, beginning with a unit vector based at r, shows that $\mathrm{Tw}([g^{-1},h^{-1}],r) > 0$, so by proposition 3.4.4 $[g,h], [g^{-1},h^{-1}] \in \mathrm{Ell}_1$. If the axes of g,h intersect in the opposite manner, similarly $\mathrm{Tw}([g^{-1},h^{-1}],r) < 0$. We will treat the case shown, i.e. $[g,h] \in \mathrm{Ell}_1$; if $[g,h] \in \mathrm{Ell}_{-1}$ then all these arguments are mirror reversed.

From proposition 3.4.4 we have $\mathrm{Tw}([g^{-1},h^{-1}],r) \in (0,2\pi)$. We will treat the cases $\mathrm{Tw}([g^{-1},h^{-1}],r) \in (0,\pi]$ and $\mathrm{Tw}([g^{-1},h^{-1}],r) \in [\pi,2\pi)$ separately.

First assume $\mathrm{Tw}([g^{-1},h^{-1}],r) \in [\pi,2\pi)$, and consider $C_\epsilon(r)$ for ϵ small. In a picture such as 7.7, with r at the top of the picture and the other fixed points below, then r can be chosen so that $p, [g^{-1},h^{-1}]p$ both lie above r and $[g^{-1},h^{-1}]p$ lies to the right of p, as shown in figure 7.8. (Note we make no claims yet about whether $\mathcal{P}(g,h,p)$ is non-degenerate or simple; in figure 7.8 it self-intersects.) Let $\varphi = 2\pi - \mathrm{Tw}([g^{-1},h^{-1}],q)$, so that $\varphi \in (0,\pi]$ and the angle $\angle pr([g^{-1},h^{-1}]p) = \varphi$, as shown in figure 7.8. Let α denote the angle $\angle(hr)r(h^{-1}ghr)$. As p moves around $C_\epsilon(r)$, its images move around $C_\epsilon(hq), C_\epsilon(ghq), C_\epsilon(h^{-1}ghq)$ with the same angular velocity.

We rotate p in the diagram shown along $C_\epsilon(r)$, to the point p_0 lying $\frac{\pi-\varphi}{2}$ past the point where $C_\epsilon(r)$ intersects the geodesic segment $r \longrightarrow hr$, as in figure 7.8. It follows that p_0 and $[g^{-1},h^{-1}]p_0$ both lie the same perpendicular (hyperbolic) distance

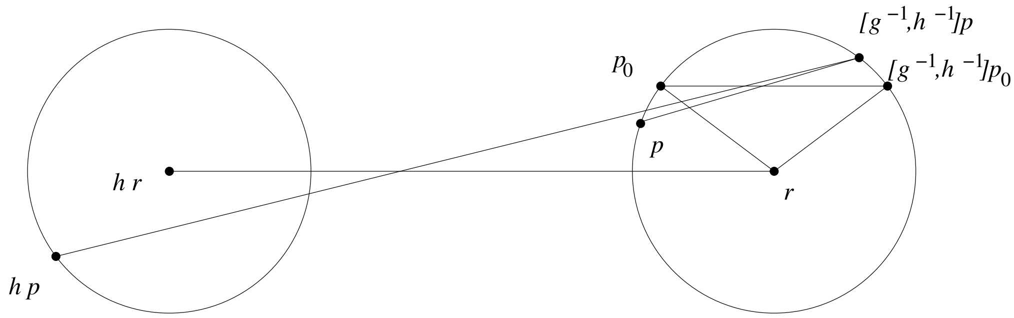

Figure 7.9: Points in $\mathcal{P}(g,h,p)$ relative to the line $hr \longrightarrow r$

from the line through r and hr.

We claim that, while for this p_0 the pentagon $\mathcal{P}(g,h,p_0)$ is not simple, for any p lying anticlockwise of and close to p_0, we may take ϵ sufficiently small so that $\mathcal{P}(g,h,p)$ is simple.

It is clear, first of all, that for any such p the segments $h^{-1}ghp \longrightarrow p$ and $p \longrightarrow [g^{-1},h^{-1}]p$ intersect only at p. The sides of the pentagon not shown in figure 7.8 clearly pose no problem. Thus to show $\mathcal{P}(g,h,p)$ is simple it is sufficient to show that the segment $[g^{-1},h^{-1}]p \longrightarrow hp$ does not intersect $h^{-1}ghp \longrightarrow p$ and intersects $[g^{-1},h^{-1}]p \longrightarrow p$ only at $[g^{-1},h^{-1}]p$.

Consider the heights of various points (i.e. Fermi coordinates) with respect to the line $r \longrightarrow hr$. It is sufficient to show that, in the arrangement of figure 7.9, the segment $hp \longrightarrow [g^{-1},h^{-1}]p$ lies above the segment $p \longrightarrow [g^{-1},h^{-1}]p$.

Now for p anticlockwise of p_0, we see that p is lower than $[g^{-1},h^{-1}]p$ with respect to the line $r \longrightarrow hr$. But by taking ϵ sufficiently small, the segment $[g^{-1},h^{-1}]p \longrightarrow hp$ can be made arbitrarily flat, since it rises by a height at most 2ϵ over some fixed distance. As this segment becomes sufficiently flat, it will lie above the segment $p \longrightarrow [g^{-1},h^{-1}]p$ as required.

By a similar argument, we may rotate p anticlockwise until $[g^{-1},h^{-1}]p$ lies $\frac{\pi-\varphi}{2}$ anticlockwise past the intersection of $C_\epsilon(r)$ with the segment $r \longrightarrow h^{-1}ghr$. While $\mathcal{P}(g,h,p)$ is not simple for this $p = p_0$, for any p up to this point, we may take ϵ sufficiently small so that $\mathcal{P}(g,h,p)$ is simple.

Thus we have found an open arc of angle $\pi+\alpha$ of directions from r, and for each direction there exists ϵ such that perturbing p in this direction by a distance less than ϵ gives $\mathcal{P}(g,h,p)$ non-degenerate and simple. This is shown in figure 7.10.

In particular, there is a *closed* arc of angle π in which p may be chosen such that $\mathcal{P}(g,h,p)$ is simple; and then by compactness we may choose an ϵ uniformly.

Now we consider $\text{Tw}([g^{-1},h^{-1}],r) \in (0,\pi]$. Let $\varphi = \text{Tw}([g^{-1},h^{-1}],r)$, so $\varphi \in (0,\pi]$. Again we rotate p around $C_\epsilon(r)$ and inquire as to when $\mathcal{P}(g,h,p)$ remains

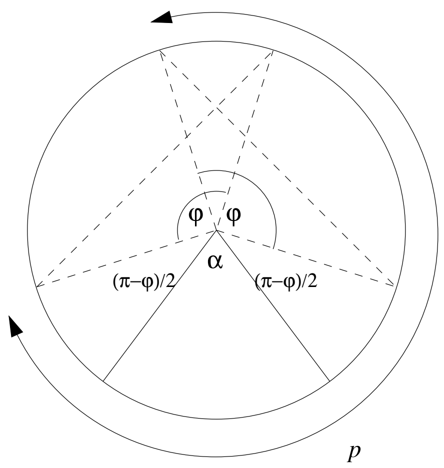

Figure 7.10: Directions p may be perturbed when $\mathrm{Tw}([g^{-1},h^{-1}],r) \in [\pi, 2\pi)$.

simple. We rotate p to the point p_0 where $[g^{-1},h^{-1}]p_0$ lies $\frac{\pi-\varphi}{2}$ clockwise of the intersection of $C_\epsilon(r)$ with the segment $r \longrightarrow h^{-1}ghr$. See figure 7.11.

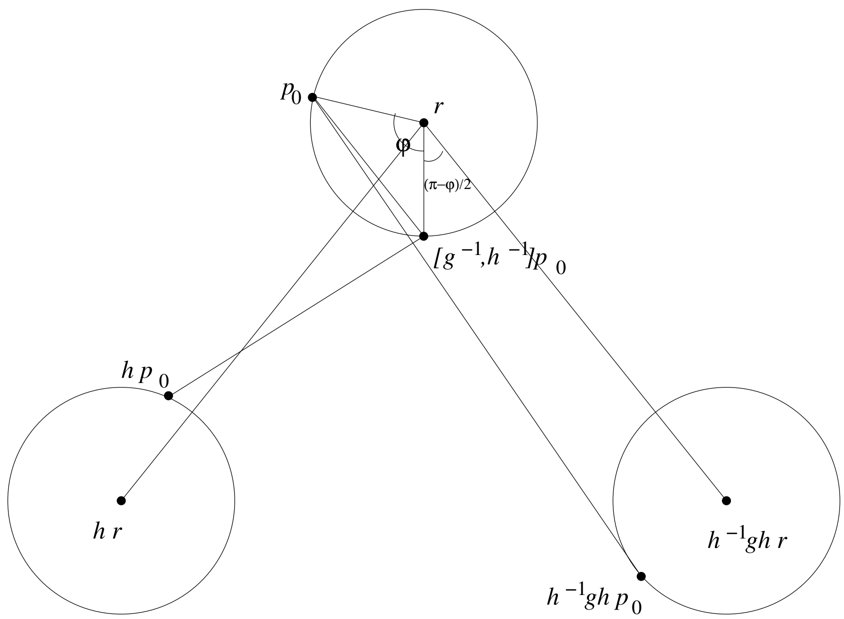

Figure 7.11: The situation of $\mathcal{P}(g,h,p)$ in the case $\mathrm{Tw}([g^{-1},h^{-1}],r) \in (0,\pi]$.

Consider heights of points with respect to the line $r \longrightarrow h^{-1}ghq$. We see that for p clockwise of p_0, the point p lies closer to the line than $[g^{-1},h^{-1}]p$. As before, taking ϵ sufficiently small we may make the segment $p \longrightarrow h^{-1}ghp$ arbitrarily flat so that $\mathcal{P}(g,h,p)$ remains simple.

Similarly, we may rotate p clockwise to the point p_1 lying $\frac{\pi-\varphi}{2}$ anticlockwise of the intersection point of $C_\epsilon(r)$ and the segment $r \longrightarrow hr$. For all p clockwise of p_0 and anticlockwise of p_1 the pentagon $\mathcal{P}(g,h,p)$ remains simple. Thus we obtain

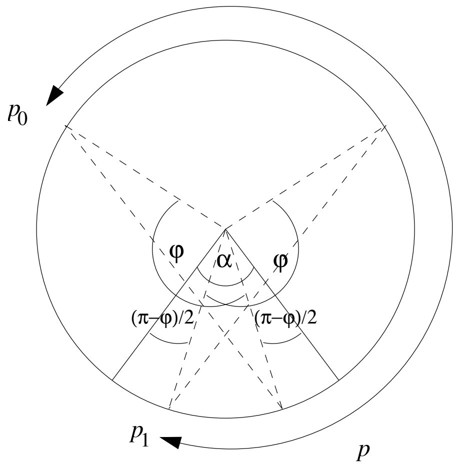

Figure 7.12: Directions p may be perturbed when $\text{Tw}([g^{-1}, h^{-1}], r) \in (0, \pi]$.

another open arc of angle $\pi + \alpha$ such that for any p in the interior of this arc, ϵ can be chosen sufficiently small that $\mathcal{P}(g, h, p)$ remains simple. Again there is a closed arc of angle π, and a uniform ϵ, giving good choices for p.

Thus, for p chosen as described above, in either case, $\mathcal{P}(g, h, p)$ is non-degenerate and bounds an embedded disc. By lemma 5.1.5, ρ is the holonomy of a hyperbolic cone-manifold structure on S with no interior cone points and at most one corner point. There is some freedom in the choice of p, and there is a closed semicircular disc of radius ϵ with centre r in which p may be chosen arbitrarily (except that $p \neq r$!); but for p chosen far from r there is no reason to expect $\mathcal{P}(g, h, p)$ to bound an immersed disc. The basis G, H can be chosen arbitrarily.

We now consider the corner angles obtained. We have considered the two cases $\text{Tw}([g^{-1}, h^{-1}], r) \in [\pi, 2\pi)$ and $\text{Tw}([g^{-1}, h^{-1}], r) \in (0, \pi]$ separately. Examining figures 7.8 and 7.11 we see that in both cases, the segment $p \to [g^{-1}, h^{-1}]p$ bounds $\mathcal{P}(g, h, p)$ on its left; so ∂S traversed in the direction of $[G, H]$ bounds S on its left. By lemma 5.4.1, the corner angle $\theta \equiv \pi - \text{Tw}([g^{-1}, h^{-1}], p)$ modulo 2π. On the other hand, $\mathcal{P}(g, h, p)$ is obtained by perturbing the situation when $p = r$, which gives a quadrilateral, hence with area less than 2π; so for p sufficiently close to r, $\mathcal{P}(g, h, p)$ has area in $(0, 2\pi)$, and $\theta \in (\pi, 3\pi)$.

In the case $\text{Tw}([g^{-1}, h^{-1}], r) \in [\pi, 2\pi)$, we have $\text{Tw}([g^{-1}, h^{-1}], p) \in [\pi, 2\pi)$ also, by lemma 3.4.2, in fact $\text{Tw}([g^{-1}, h^{-1}], x) \in [\pi, 2\pi)$ for all $x \in \mathbb{H}^2$, so by proposition 3.6.3, $\Theta([g^{-1}, h^{-1}]) \in [\pi/2, \pi)$. We have $\theta \in (-\pi, 0]$ mod 2π, and hence $\theta \in (\pi, 2\pi]$. Therefore, $\theta = 3\pi - \text{Tw}([g^{-1}, h^{-1}], p)$. At least for p moving within a small semi-circular disc centred at r, θ varies continuously, and so $\theta \in (\pi, 2\pi]$ throughout this region.

In the case $\text{Tw}([g^{-1}, h^{-1}], r) \in (0, \pi]$ we have $\text{Tw}([g^{-1}, h^{-1}], p) \in (0, \pi]$ again by lemma 3.4.2, and $\Theta([g^{-1},h^{-1}]) \in (0,\pi/2]$, and $\theta \in [0,\pi)$ mod 2π, so $\theta \in [2\pi, 3\pi)$. Again, for p in a small semicircular disc around r, $\theta \in [2\pi, 3\pi)$ also. Therefore $\theta = 3\pi - \mathrm{Tw}([g^{-1},h^{-1}],p)$ in this case as well. This justifies our description in chapter 5 of the "small angle" and "large angle" cases.

If $[g,h],[g^{-1},h^{-1}] \in \mathrm{Ell}_{-1}$ then clearly the same arguments go through. We obtain a hyperbolic cone-manifold structure on S, with the orientation reversed: $p \to [g^{-1},h^{-1}]p$ bounds $\mathcal{P}(g,h,p)$ on its right, and ∂S traversed in the direction of $[G,H]$ bounds S on its right. If $\mathrm{Tw}([g^{-1},h^{-1}],r) \in (-2\pi,-\pi]$ then we again have a "small angle" case and $\theta \in (\pi, 2\pi]$, with $\theta = 3\pi + \mathrm{Tw}([g^{-1},h^{-1}],p)$. If $\mathrm{Tw}([g^{-1},h^{-1}],r) \in [-\pi, 0)$ then we have a "large angle" case and $\theta \in [2\pi,3\pi)$, $\theta = 3\pi + \mathrm{Tw}([g^{-1},h^{-1}],p)$ also.

We record our conclusions.

Proposition 7.4.1 *Let $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ be a representation with $\mathrm{Tr}[g,h] \in (-2,2)$. Suppose $[g,h] \in \mathrm{Ell}_1$ (resp. Ell_{-1}). Let r denote the fixed point of $[g,h]$. Then there exists a closed semicircular disc $D_\epsilon(q)$ with centre r such that if p is chosen anywhere in this disc, except r, then $\mathcal{P}(g,h,p)$ is simple and non-degenerate, giving a hyperbolic cone-manifold structure on S with no cone points and one corner point of angle θ. The boundary ∂S, traversed in the direction $[G,H]$ bounds S on its left (resp. right). The corner angle $\theta = 3\pi - \mathrm{Tw}([g^{-1},h^{-1}],p)$ (resp $3\pi + \mathrm{Tw}([g^{-1},h^{-1}],p)$).* ∎

In some sense, the cone-manifold structures on S can be thought of as "pushing out" or "extending" the original cone point r on the hyperbolic torus obtained by taking $p = r$. This is in some way analogous to the extension or truncation of a complete hyperbolic structure in the case $\mathrm{Tr}[g,h] < -2$, and the truncation of a cusped hyperbolic structure in the case $\mathrm{Tr}[g,h] = -2$. For p sufficiently close to r, the axes of g and h project to geodesics in the free homotopy classes of G and H.

This concludes our discussion of the case $\mathrm{Tr}[g,h] \in (-2,2)$.

7.5 The case $\mathrm{Tr}[g,h]=2$: reducible representations

We have shown in lemma 6.1.3 that ρ is reducible precisely when $\mathrm{Tr}[g,h] = 2$. Thus abelian representations are reducible. But by lemma 6.5.1 reducible virtually abelian representations are abelian. So will show that the abelian representations do not give cone-manifold structures of the desired type; and we will show that the reducible non-abelian (hence not virtually abelian) representations do give cone-manifold structures of the desired type.

Lemma 7.5.1 *An abelian representation is not the holonomy of any hyperbolic cone manifold structure on S with no interior cone points and at most one corner point.*

PROOF Let ρ be abelian. So for any basis G, H of $\pi_1(S)$ (with basepoint on ∂S), g, h commute. Hence for any $p \in \mathbb{H}^2$, $p = [g^{-1}, h^{-1}]p$, so $\mathcal{P}(g, h, p)$ has a degenerate boundary edge. So by lemma 5.1.5, ρ is not the holonomy of any such cone-manifold structure. ∎

Let us now construct a hyperbolic cone-manifold structure when ρ is non-abelian and reducible. Lemma 6.5.2 describes the situation in this case.

So first assume that one of g, h is parabolic and the other hyperbolic, with a common fixed point. After possibly reordering our basis and replacing G with its inverse, we may conjugate and assume that in the upper half-plane model $g(z) = z+1$ and $h(z) = e(z - f)$, for some $e > 0$ and some $f \in \mathbb{R}$. Since h is hyperbolic, $e \neq 1$. Take $p \in \mathbb{H}^2$ and let $p = x + iy = (x, y)$; we will compute the coordinates of the vertices of $\mathcal{P}(g, h, p)$.

$$\begin{aligned} p &= (x, y) \\ hp &= (e(x-f), ey) \\ ghp &= (e(x-f)+1, ey) \\ h^{-1}ghp &= \left(x + \frac{1}{e}, y\right) \\ [g^{-1}, h^{-1}]p &= \left(x + \frac{1}{e} - 1, y\right) \end{aligned}$$

Thus, respectively as $e \in (1, \infty)$ or $(0, 1)$, we obtain the situations of figures 7.13 or 7.14. In both cases, we see that for any choice of $p \in \mathbb{H}^2$, the pentagon $\mathcal{P}(g, h, p)$ is non-degenerate and bounds an embedded disc. So by lemma 5.1.5, ρ is the holonomy of a hyperbolic cone-manifold structure on S with no interior cone point and at most one corner point.

In the first case, the segment $p \to [g^{-1}, h^{-1}]p$ bounds $\mathcal{P}(g, h, p)$ on its right; so ∂S traversed in the direction of $[G, H]$ bounds S on its right; and we see $[g^{-1}, h^{-1}]$ is parabolic, fixing ∞, translating to the left. We know $[g, h] \in \text{Par}_0$; therefore $[g, h] \in \text{Par}_0^-$. In the second case, $p \to [g^{-1}, h^{-1}]p$ bounds $\mathcal{P}(g, h, p)$ on its left; ∂S in the direction of $[G, H]$ bounds S on its left; and $[g, h] \in \text{Par}_0^+$. This is all true regardless of our choice of p.

Now assume both g, h are hyperbolic, with precisely one common fixed point, say ∞. We may assume g, h are given by

$$g : z \mapsto a(z - b), \quad h : z \mapsto e(z - f)$$

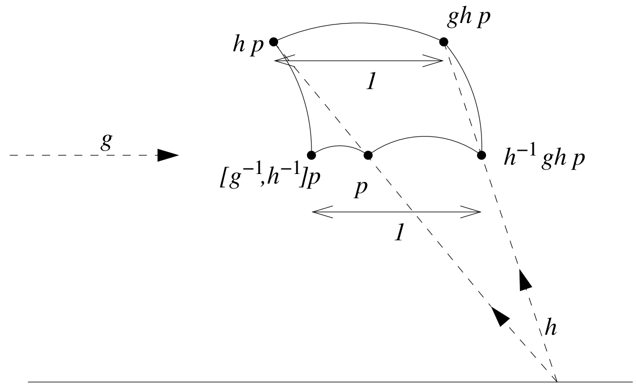

Figure 7.13: $\mathcal{P}(g,h,p)$ when $\mathrm{Tr}[g,h]=2$, non-abelian, g parabolic, and $e\in(1,\infty)$

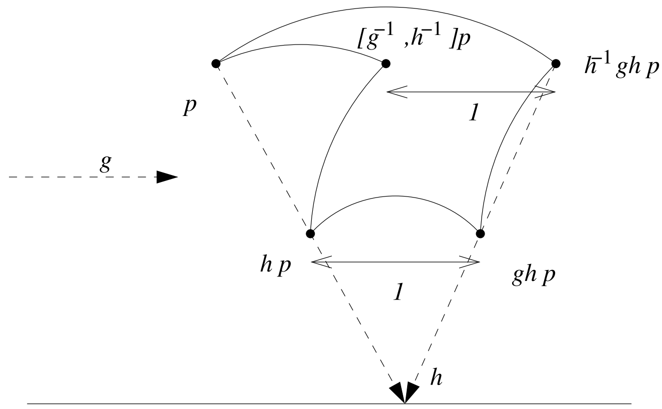

Figure 7.14: $\mathcal{P}(g,h,p)$ when $\mathrm{Tr}[g,h]=2$, non-abelian g parabolic, and $e\in(0,1)$

where $a, e > 0$ and $b, f \in \mathbb{R}$. Since g, h are hyperbolic, $a, e \neq 1$. Replacing g, h by their inverses if necessary, we may assume $a, e \in (1, \infty)$. Transposing g, h if necessary, we may assume $a < e$. Now conjugating by a parabolic of the form $z \mapsto z + c$, we may assume g, h are given by

$$g : z \mapsto az, \quad h : z \mapsto e(z - f), \quad 1 < a < e, \quad f \neq 0.$$

So the fixed points at infinity of g are $\{0, \infty\}$, and the fixed points at infinity of h are $\{\frac{ef}{e-1}, \infty\}$. We compute

$$\begin{aligned} p &= (x, y) \\ hp &= (e(x - f), ey) \\ ghp &= (ae(x - f), aey) \\ h^{-1}ghp &= (a(x - f) + f, ay) \\ [g^{-1}, h^{-1}]p &= \left(x + \frac{f}{a} - f, y \right). \end{aligned}$$

First suppose that $f < 0$. Then $\frac{ef}{e-1} < 0$, and we have a situation as in figure 7.15. Choosing p to lie above the fixed point $\frac{ef}{f-1}$, i.e. $p = (\frac{ef}{f-1}, y)$ for some y, we see that hp lies directly above p, along the (Euclidean and hyperbolic) line $\frac{ef}{e-1} \to p$. Then ghp lies above hp, along the Euclidean line $0 \to hp$; and $h^{-1}ghp$ lies below ghp, in the Euclidean segment $\frac{ef}{e-1} \to ghp$. In particular we see that the line through $\frac{ef}{e-1}, p, hp$ splits the plane with $ghp, h^{-1}ghp$ on its left, with the four (Euclidean or hyperbolic) segments $p \to hp \to ghp \to h^{-1}ghp \to p$ forming a non-degenerate simple quadrilateral. To show that $\mathcal{P}(g, h, p)$ is a non-degenerate simple pentagon it is sufficient that $[g^{-1}, h^{-1}]p$ lies on the right of the line $\frac{ef}{e-1} \to p \to hp$. But p and $[g^{-1}, h^{-1}]p$ lie at the same height, so it is sufficient that $[g^{-1}, h^{-1}]p$ lies to the right of p, i.e. $\frac{f}{a} - f > 0$. But this is true as $f < 0$ and $a > 1$. Hence $\mathcal{P}(g, h, p)$ is non-degenerate bounding an embedded disc.

Alternatively, suppose $f > 0$. Then by the mirror image of the previous argument, we find $\mathcal{P}(g, h, p)$ is non-degenerate bounding an embedded disc, again choosing p above the fixed point $\frac{ef}{e-1}$ of h. So by lemma 5.1.5, ρ is the holonomy of a hyperbolic cone-manifold structure on S with no interior cone points and at most one corner point.

In the case $f < 0$, the segment $p \to [g^{-1}, h^{-1}]p$ bounds $\mathcal{P}(g, h, p)$ on its left, so ∂S traversed homotopic to $[G, H]$ bounds S on its left. And $[g^{-1}, h^{-1}]$ is parabolic, in fact $[g, h], [g^{-1}, h^{-1}] \in \text{Par}_0^+$. In the case $f > 0$, $p \to [g^{-1}, h^{-1}]p$ bounds $\mathcal{P}(g, h, p)$ on its right; ∂S in the direction of $[G, H]$ bounds S on its right; and $[g, h] \in \text{Par}_0^-$.

Note that almost any basis is good enough to produce a g, h which works; at most we reordered the basis and replaced the basis elements with their inverses.

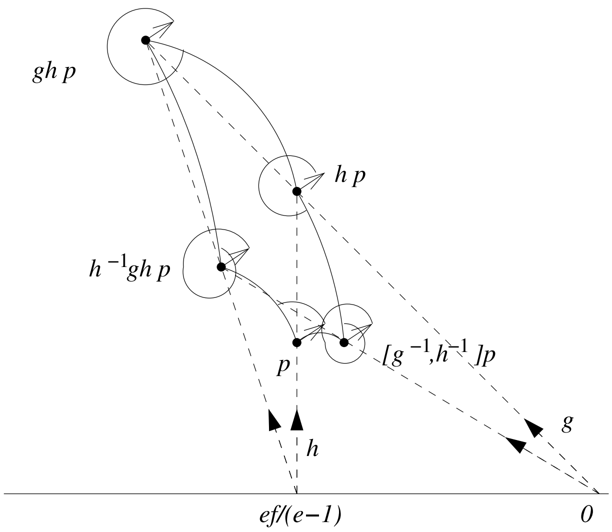

Figure 7.15: $\mathcal{P}(g, h, p)$ when $\text{Tr}[g, h] = 2$, non-abelian, g, h hyperbolic, and $f < 0$. A unit vector chase shows $\theta > 2\pi$.

And there is freedom in the choice of p also. In the first case, where one of g, h is parabolic, we can place p arbitrarily. In the second case, where both g, h are hyperbolic, we placed p arbitrarily along a certain line in $\mathbb{H}^2$. And p may certainly be perturbed from the location we chose, but in general if p is perturbed too far then the pentagon may no longer be simple. However, certainly p can be chosen arbitrarily close to the fixed point of $[g^{-1}, h^{-1}]$.

Clearly the corner angle θ varies continuously as p varies. For $[g, h] \in \text{Par}_0^+$ we have $\text{Tw}([g^{-1}, h^{-1}], p) \in (0, \pi)$ always, and $p \to [g^{-1}, h^{-1}]p$ bounding $\mathcal{P}(g, h, p)$ on its left. For $[g, h] \in \text{Par}_0^-$ we have $\text{Tw}([g^{-1}, h^{-1}], p) \in (-\pi, 0)$ and $p \to [g^{-1}, h^{-1}]p$ bounding $\mathcal{P}(g, h, p)$ on its right. Hence by lemma 5.4.1, in either case $\theta \in (0, \pi)$ or $(2\pi, 3\pi)$. Since we have determined the explicit situation, we may perform a unit vector chase in the manner of the proof of lemma 5.4.1, and obtain $\theta \in (2\pi, 3\pi)$ for p close to the point chosen above. See e.g. figure 7.15. From lemma 5.4.1 then $\theta = 3\pi - \text{Tw}([g^{-1}, h^{-1}], p)$ for the Par_0^+ case and $\theta = 3\pi + \text{Tw}([g^{-1}, h^{-1}], p)$ for the Par_0^- case. In particular, we see that the pentagon must degenerate if p wanders too far from the point chosen above; if p could venture arbitrarily far, then the pentagon could approach an ideal pentagon and we would have $\theta \to 0$, a contradiction.

Proposition 7.5.2 *Let $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ be a representation with $\text{Tr}[g, h] = 2$ for some basis G, H of $\pi_1(S)$. Then ρ is the holonomy of a hyperbolic cone manifold structure on S with no cone points and at most one corner point if and only if ρ is not virtually abelian, i.e. $[g, h] \in \text{Par}_0$. A fundamental domain for the developing map is given by $\mathcal{P}(g', h', p)$ where (G', H') is obtained from (G, H) at most by reordering* *and replacing with inverses. Suppose* $[g', h'] \in \text{Par}_0^+$ *(resp.* Par_0^-*). The point* p *may be chosen arbitrarily close to the fixed point at infinity of* $[g'^{-1}, h'^{-1}]$*. Then the boundary* ∂S*, traversed in the direction of* $[G, H]$*, bounds* S *on its left (resp. right). The corner angle* $\theta = 3\pi - \text{Tw}([g^{-1}, h^{-1}], p)$ *(resp.* $3\pi + \text{Tw}([g^{-1}, h^{-1}], p)$*).* ∎

7.6 The case $\text{Tr}[g, h] > 2$

We now come to the most difficult case. This case includes virtually abelian representations. The abelian representations all belonged to the case $\text{Tr}[g, h] = 2$; by lemma 6.4.3, the representations which are virtually abelian but not abelian are precisely those with $(\text{Tr}\, g, \text{Tr}\, h, \text{Tr}\, gh) = (x, y, z) \in V$, in the notation of section 6.4, and hence $\text{Tr}[g, h] = \kappa(x, y, z) > 2$. Our proof is in the following three subsections, which respectively prove the following three results.

Proposition 7.6.1 *Let* $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ *be a representation which is virtually abelian but not abelian. Then* ρ *is not the holonomy of any hyperbolic cone manifold structure on* S *with no interior cone points and at most one corner point.*

Proposition 7.6.2 *Let* G, H *be a basis of* $\pi_1(S)$ *and let* $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ *be a representation with* $\text{Tr}[g, h] > 2$ *which is not virtually abelian. Then there exists a basis* G', H' *of* $\pi_1(S)$ *such that*

$$(x, y, z) = (\text{Tr}\, g', \text{Tr}\, h', \text{Tr}\, g'h') \in (2, \infty)^3.$$

Proposition 7.6.3 *Let* $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ *be a representation which is not virtually abelian, and suppose there exists a basis* G, H *of* $\pi_1(S)$ *such that* $\text{Tr}[g, h] > 2$ *and* $(x, y, z) = (\text{Tr}\, g, \text{Tr}\, h, \text{Tr}\, gh) \in (2, \infty)^3$*. Then* ρ *is the holonomy of a hyperbolic cone-manifold structure on* S *with no interior cone points and at most one corner point.*

7.6.1 Virtually abelian degeneration

We now prove proposition 7.6.1, so let ρ be a representation which is virtually abelian but not abelian. From section 6.4, ρ has character in V, and sends two of g, h, gh to half turns about two distinct points, and the other to a translation along the line connecting those points.

Lemma 7.6.4 *Let* $g, h \in PSL_2\mathbb{R}$ *such that* $(\text{Tr}\, g, \text{Tr}\, h, \text{Tr}\, gh) \in V$*. Then for any* $p \in \mathbb{H}^2$*, the pentagon* $\mathcal{P}(g, h, p)$ *does not bound an immersed open disc in* $\mathbb{H}^2$*.*

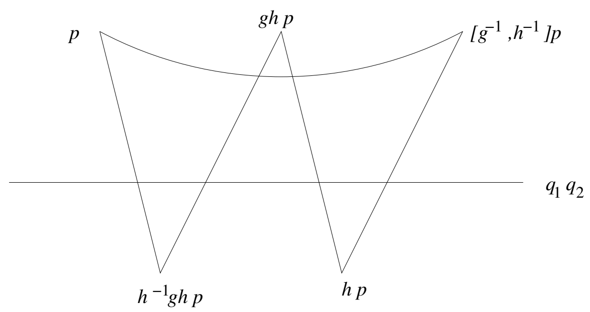

Figure 7.16: $\mathcal{P}(g,h,p)$ does not bound an immersed disc in case (i).

PROOF Two of $\{g,h,gh\}$ are half-turns about distinct points $q_1,q_2 \in \mathbb{H}^2$, and the third is hyperbolic with axis q_1q_2. There are three possible cases.

(i) g,h are half-turns, gh is hyperbolic.

(ii) h,gh are half-turns, g is hyperbolic.

(iii) g,gh are half-turns, h is hyperbolic.

In each case, all of g,h,gh preserve the line q_1q_2. Thus $[g,h]$ preserves the line q_1q_2, and its orientation. As $\mathrm{Tr}[g,h] > 2$, $[g,h]$ is hyperbolic with axis q_1q_2. We consider the three cases separately.

Case (i). If $p \in q_1q_2$ then all vertices of $\mathcal{P}(g,h,p)$ lie on q_1q_2 and the pentagon is clearly degenerate, so clearly does not bound an immersed disc. Otherwise let the perpendicular distance from p to q_1q_2 be d. Consider Fermi coordinates on $\mathbb{H}^2$ with axis q_1q_2, and let $p=(\alpha,d)$. We take coordinates so that g acts as $(y,z)\mapsto(-y+a,-z)$ and h acts as $(y,z)\mapsto(-y,-z)$, for some nonzero $a\in\mathbb{R}$. Then we have

$$\begin{aligned} p &= (\alpha,d),\\ hp &= (-\alpha,-d),\\ ghp &= (\alpha+a,d),\\ h^{-1}ghp &= (-\alpha-a,-d),\\ [g^{-1},h^{-1}]p &= (\alpha+2a,d). \end{aligned}$$

Note that, regardless of the signs of α and a, the point ghp lies between p and $[g^{-1},h^{-1}]p$ on the curve at height d from q_1q_2. Since $a\neq 0$ these three points are distinct. But now ghp lies on the opposite side of the geodesic segment $p\longrightarrow[g^{-1},h^{-1}]p$ to the points hp and ghp. It follows that $\mathcal{P}(g,h,p)$ does not bound an immersed disc. See figure 7.16.

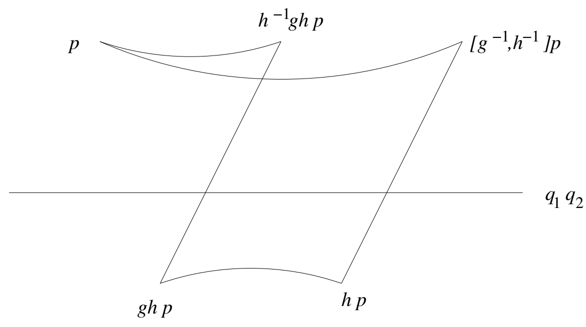

Figure 7.17: $\mathcal{P}(g,h,p)$ does not bound an immersed disc in case (ii).

Case (ii). Again take Fermi coordinates with axis l. We may assume that g acts as $(y,z)\mapsto(y+c,z)$ for some $c\neq 0$, and that h acts as $(y,z)\mapsto(-y,-z)$. Let $p=(\alpha,d)$. If $d=0$ then all the vertices of $\mathcal{P}(g,h,p)$ lie on l so the pentagon clearly does not bound an immersed disc. Otherwise we have

$$\begin{aligned} p &= (\alpha,d),\\ hp &= (-\alpha,-d),\\ ghp &= (-\alpha+c,-d),\\ h^{-1}ghp &= (\alpha-c,d),\\ [g^{-1},h^{-1}]p &= (\alpha-2c,d). \end{aligned}$$

Now $h^{-1}ghp$ lies between p and $[g^{-1},h^{-1}]p$ on the curve at height d from l. But now $h^{-1}ghp$ lies on the opposite side of the geodesic segment $p\longrightarrow[g^{-1},h^{-1}]p$ to the points hp and ghp. Again $\mathcal{P}(g,h,p)$ cannot bound an immersed disc. See figure 7.17.

Case (iii). This is similar to case (ii). ∎

Proof (Of proposition 7.6.1) Recall that ρ is a holonomy representation of a hyperbolic cone manifold structure on S with no interior cone points and at most one corner point if and only if there exist a basis G,H of $\pi_1(S)$ and a point $p\in\mathbb{H}^2$ such that $\mathcal{P}(g,h,p)$ is non-degenerate and bounds an immersed disc, by lemma 5.1.5. Since ρ is virtually abelian but not abelian, for any basis G,H, $(x,y,z)\in V$, by lemma 6.4.2. Then by the previous lemma, for all p, $\mathcal{P}(g,h,p)$ does not bound an immersed open disc. So ρ cannot be such a holonomy representation. ∎

7.6.2 An algorithm to increase traces

By theorem 6.1.2, any triple (x,y,z) satisfying $\kappa(x,y,z)>2$, lies in $X(S)$, the character variety, i.e. $(x,y,z)=(\operatorname{Tr}g,\operatorname{Tr}h,\operatorname{Tr}gh)$ for some representation ρ. The subset of $X(S)$ with $\mathrm{Tr}[g, h] = \kappa(x, y, z) > 2$ and corresponding to virtually abelian representations is

$$V = \left\{ (x, y, z) \in \mathbb{R}^3 : \begin{array}{c} \text{two coordinates are zero and} \\ \text{the third is} > 2 \text{ in magnitude} \end{array} \right\}.$$

We have fully investigated the effect of changes of basis on characters $(x, y, z) = (\mathrm{Tr}\, g, \mathrm{Tr}\, h, \mathrm{Tr}\, gh)$ in section 4.5. In light of proposition 6.3.2, proposition 7.6.2 is reduced to the following purely algebraic claim.

Lemma 7.6.5 *Let $(x, y, z) \in \mathbb{R}^3$ satisfy $x^2 + y^2 + z^2 - xyz > 4$ and $(x, y, z) \notin V$. Then under the equivalence relation generated by permutations of coordinates and*

$$(x, y, z) \sim (x, y, xy - z), \quad (x, y, z) \sim (-x, -y, z),$$

we have $(x, y, z) \sim (x', y', z')$ for some $(x', y', z') \in (2, \infty)^3$.

We will give an algorithm to obtain such an (x', y', z'). We will consider the following subsets of $\mathbb{R}^3$, each to be treated separately.

$$\begin{aligned} R_1 &= (2, \infty) \times (2, \infty) \times (2, \infty) \\ R_2 &= (-\infty, -2) \times (2, \infty) \times (2, \infty) \\ R_3 &= [-2, 2] \times (2, \infty) \times (2, \infty) \\ R_4 &= [0, 2] \times [0, 2] \times (2, \infty) \\ R_5 &= [-2, 0] \times [0, 2] \times (2, \infty) \\ R_6 &= [0, 2] \times [0, 2] \times [0, 2] \\ R_7 &= [-2, 0] \times [0, 2] \times [0, 2] \end{aligned}$$

Since sign changes on two coordinates and permutations of coordinates are valid moves, we may reorder (x, y, z) so that $|x| \le |y| \le |z|$; and then change signs until $y, z \ge 0$. This point lies in some R_i. Thus every point in $\mathbb{R}^3$ is equivalent to a point in $\cup R_i$. We will show that every point in R_i, for $2 \le i \le 7$, is equivalent to a point in some R_j for $j < i$. It will follow that every point in $\mathbb{R}^3$ is equivalent to a point in R_1 as required.

We will always proceed by a greedy algorithm: permute coordinates so that $x \le y \le z$ and then apply the Markoff move $(x, y, z) \mapsto (yz - x, y, z)$. So it is worth examining this algebra first.

Recall that

$$\kappa(x, y, z) = \mathrm{Tr}[g, h] = x^2 + y^2 + z^2 - xyz - 2$$

where κ is invariant under any automorphism of the free group; in particular under a change of basis $(g, h) \mapsto (g^{-1}, h)$. Letting $x' = yz - x$ be the number replacing x after the Markoff move is applied, we see that x, x' are the roots of the quadratic

$$t^2 - yzt + y^2 + z^2 - \kappa - 2 = 0$$

where $\kappa > 2$ is a constant. Here we think of y, z as constants. The quadratic has discriminant

$$\Delta = (y^2 - 4)(z^2 - 4) + 4\kappa - 8$$

and roots given by

$$x, x' = \frac{yz \pm \sqrt{\Delta}}{2},$$

and turning point at $t = yz/2$. We now turn to each of the regions R_2 through to R_7 in turn.

The region R_7

After possibly reordering coordinates we may assume

$$-2 \le x \le 0 \le y \le z \le 2.$$

We now simply take

$$(x', y', z') = (yz - x, y, z)$$

in which all coordinates are non-negative, so that (x', y', z') (after reordering coordinates) lies in R_4 or R_6.

The region R_6

This is the most difficult case. After possibly reordering coordinates we may assume

$$0 \le x \le y \le z \le 2.$$

We need a rather technical lemma; the analogous inequality will be clear in other cases.

Lemma 7.6.6 *Suppose* $0 \le x \le y \le z \le 2$, *and* $x^2 + y^2 + z^2 - xyz > 4$. *Then* $yz - x > y$.

PROOF We minimize the function $f(x, y, z) = yz - x - y$ subject to the constraints

$$0 \le x \le y \le z \le 2, \quad x^2 + y^2 + z^2 - xyz \ge 4.$$

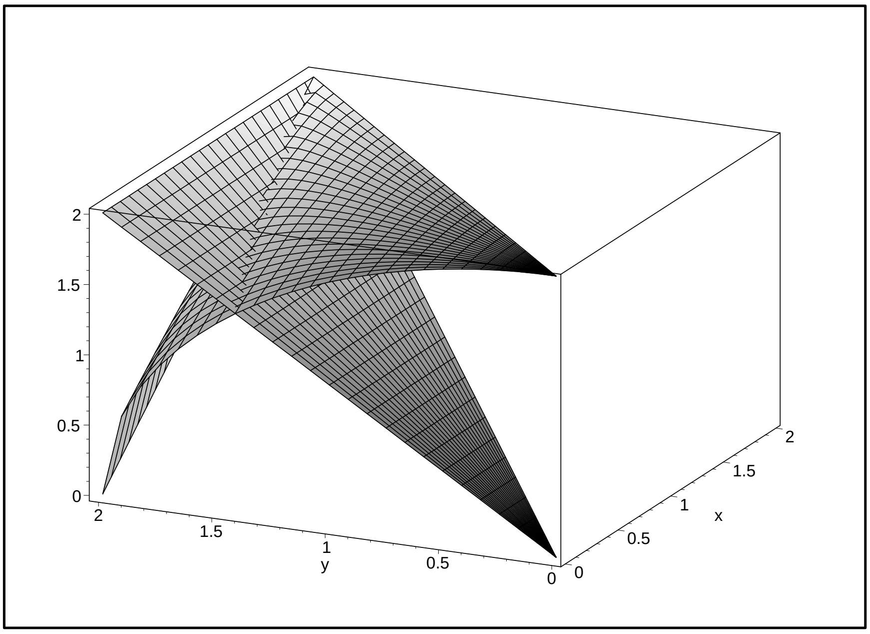

Figure 7.18: The region D. The bounding surfaces $x^2+y^2+z^2-xyz=4$ and $y=z$ are shown. The surfaces $x=0, z=2$ are sides of the bounding box; the surface $x=y$ is implied by the side of the plotted surfaces. The region D is the region lying above $y=z$ and $x^2+y^2+z^2-xyz=4$.

The region $D \subset \mathbb{R}^3$ defined by these inequalities is connected, compact, and bounded by the surfaces $x=0$, $x=y$, $y=z$, $z=2$, $x^2+y^2+z^2-xyz=4$. (Actually we will see that only four of these suffice to bound D; the surface $x=y$ is not necessary.) Four of these surfaces are obviously planes; the last is connected and forms a single face of the region D, whose edges are curves along the other planes. It is depicted in figure 7.18.

We have

$$\nabla f = (-1, z-1, y)$$

which is never zero. Thus any minimum is achieved on ∂D. We claim this minimum is zero and is only achieved at points where $x^2+y^2+z^2-xyz=4$. It will follow that when $x^2+y^2+z^2-xyz>4$ we have $f(x,y,z)=yz-x-y>0$ strictly, proving the result.

Note that if $z \leq 1$ then $x, y \leq 1$ also so that

$$x^2+y^2+z^2-xyz \leq x^2+y^2+z^2 \leq 3,$$

contradicting $x^2+y^2+z^2-xyz \geq 4$. Thus $z>1$.

On the surface $x=0$ we have $f(0,y,z)=yz-y=y(z-1)$, which is non-negative by our previous remark. Since $z>1$ it takes the value 0 only when $y=0$. Then

$x^2 + y^2 + z^2 - xyz = z^2 \geq 4$ but $z \leq 2$ so $z = 2$. Thus f achieves a minimum of 0 at $(0, 0, 2)$.

The surface $x = y$ actually only intersects ∂D in the line $x = y, z = 2$, which also lies on the surfaces $x^2+y^2+z^2-xyz = 4$ and $z = 2$. If $x = y$ then $x^2+y^2+z^2-xyz \geq 4$ implies $2x^2 + z^2 - x^2z \geq 4$, which simplifies to $(z - x^2 + 2)(z - 2) \geq 0$. Suppose $z \neq 2$, then $z \leq x^2 - 2$. Thus $x \leq z \leq x^2 - 2$, but for $0 \leq x \leq 2$ we have $x^2 - 2 \leq x$ as $x^2 - x - 2 = (x - 2)(x + 1) \leq 0$. Hence $x^2 - 2 = x$ and $x = 2$, giving the sole point $(2, 2, 2)$. Thus $x = y$ intersects ∂D precisely on the line segment $x = y, z = 2$, along which $f(x, x, 2) = 0$. So f achieves a minimum along this segment.

The surface $y = z$ gives us $f(x, y, y) = y^2 - x - y$, which we minimize given the constraints $0 \leq x \leq y \leq 2$ and $x^2 + 2y^2 - xy^2 \geq 4$, which simplifies to $(x - y^2 + 2)(x - 2) \geq 0$. So $x = 2$ or $y^2 - x \geq 2$. If $x = 2$ we have $y, z = 2$ also and obtain $f(2, 2, 2) = 0$. Otherwise $y^2 - x \geq 2$, thus $f(x, y, y) = y^2 - x - y \geq 2 - y \geq 0$ with equality only if $y = 2$. Along this line $y = z = 2$ we have $f(x, 2, 2) = 2 - x$, so f achieves a minimum of 0 at $(2, 2, 2)$.

Along the plane $z = 2$ we have $f(x, y, 2) = y - x$ which is clearly non-negative, with a minimum of 0 if and only if $x = y$. Along this line $x = y$, $z = 2$ we have $f(x, x, 2) = 0$. Thus f achieves a minimum of 0 along the segment $\{(x, x, 2) : 0 \leq x \leq 2\}$.

On the surface $T = \{(x, y, z) : x^2 + y^2 + z^2 - xyz = 4, 0 \leq x \leq y \leq z \leq 2\}$ we note than we have already checked ∂T, since every point of ∂T lies in some other face of ∂D also. Thus we now need only check $f \geq 0$ throughout the interior of T. We use Lagrange multipliers, so set $g(x, y, z) = x^2 + y^2 + z^2 - xyz - 4$. We minimize $f(x, y, z)$ subject to the constraint $g = 0$. This occurs when $\nabla f = \lambda \nabla g$ for some $\lambda \in \mathbb{R}$, giving us

$$(-1, z - 1, y) = \lambda \left(2x - yz, 2y - xz, 2z - xy\right).$$

Now if $y = 0$ then also $x = 0$, which we already know about, and is not on the interior of T. And if $2z - xy = 0$ then $xy = 2z$ despite the inequalities $x \leq z$ and $y \leq 2$. Thus either $x = y = z = 2$ or some coordinate is zero, hence $x = 0$. These points we have already seen. So we may divide the second by the third coordinates to obtain

$$\frac{z - 1}{y} = \frac{2y - xz}{2z - xy},$$

which simplifies to

$$2z(z - 2) = y(2y - x).$$

But this equality is highly suspicious as $2z(z - 2) \leq 0$ and $y(2y - x) \geq 0$. Thus $z \in \{0, 2\}$ and $y \in \{0, x/2\}$. If $z = 0$ then $x = y = 0$, but $g(0, 0, 0) = -4 \neq 0$ so this point is not in T. The case $z = 2$ we have already dealt with. If $y = 0$ then $x = 0$ also, with which we have dealt. And if $y = x/2$ then $0 \leq x \leq y = x/2$ which implies $x = y = 0$, disposed with now for the final time.

Thus f achieves a minimum of 0, precisely at the points $(x, x, 2)$ for $0 \leq x \leq 2$. We see that $x^2 + y^2 + z^2 - xyz = 4$ for all these points, as required. ■

We now apply our greedy algorithm. The sequence (x_n, y_n, z_n) is defined inductively by setting (x_0, y_0, z_0) equal to our given triple (x, y, z) and letting $(x_{n+1}, y_{n+1}, z_{n+1})$ be the triple obtained by taking

$$\{y_n z_n - x_n, y_n, z_n\}$$

and reordering so $x_{n+1} \leq y_{n+1} \leq z_{n+1}$.

Now the lemma tells us that $y_n z_n - x_n \geq y_n$, so that all coordinates remain non-negative. At most one of (x, y, z) can be zero: if two are zero then ρ is virtually abelian; if three are zero we have a contradiction to $\kappa(x, y, z) > 2$. The first application of the Markoff move makes all coordinates positive, after which they remain positive and non-decreasing. The sum $x_n + y_n + z_n$ is strictly increasing, and in fact

$$\begin{aligned}(x_{n+1} + y_{n+1} + z_{n+1}) - (x_n + y_n + z_n) &= y_n z_n - 2x_n \\ &= x_n' - x_n \\ &= \sqrt{(y_n^2 - 4)(z_n^2 - 4) + 4\kappa - 8} \\ &\geq 2\sqrt{\kappa - 2}.\end{aligned}$$

This last inequality follows since, if one of x_n, y_n, z_n becomes larger than 2 then our point lies in a different region (namely R_4) and we have completed the argument. Otherwise $y_n^2 - 4, z_n^2 - 4 \leq 0$ and their product is non-negative.

Thus the sum $x_n + y_n + z_n$ increases each iteration by at least $2\sqrt{\kappa - 2} > 0$. It follows that after a finite number of steps this sum becomes larger than 6, and hence one of the coordinates becomes larger than 2, moving our point into R_4.

The region R_5

Here $-2 \leq x \leq 0 \leq y \leq 2 < z$. We take $(x', y', z') = (yz - x, y, z)$. Now all coordinates are non-negative and $z > 2$ so that (x', y', z'), after permuting coordinates to put them in ascending order, lies in R_4 or R_3.

The region R_4

After possibly permuting coordinates we may assume that $0 \leq x \leq y \leq 2 < z$. We will apply the greedy algorithm. So let (x_0, y_0, z_0) be the given (x, y, z) and for

$n \geq 0$ inductively let $(x_{n+1}, y_{n+1}, z_{n+1})$ be the triple obtained by taking

$$(y_n z_n - x_n, y_n, z_n)$$

and permuting coordinates to put them in ascending order.

Upon applying the move $(x, y, z) \mapsto (yz - x, y, z)$ we see that

$$yz - x \geq 2y - x \geq y$$

so that all coordinates remain non-negative at each stage, at least one coordinate is greater than 2, and the minimum of the three coordinates is non-decreasing. If two coordinates become greater than 2 then we are in the region R_3. The minimum of the three coordinates is strictly increasing if and only if $y \neq 0$.

The case $y = 0$ corresponds to triples of the form $(0, 0, z)$, and the condition $\kappa > 2$ implies $z > 2$. This corresponds to a virtually abelian representation and $(x, y, z) \in V$ in this case. Any equivalent (x', y', z') lies in V also.

In any other case we have

$$0 \leq x \leq y \left\{ \begin{array}{cccc} \leq & 2 & < & z \\ < & yz - x & = & x' \end{array} \right.$$

We show that a finite number of these moves suffices to make two coordinates greater than 2.

We have $x < x'$ as the two roots of our quadratic, so

$$x = \frac{yz - \sqrt{(y^2-4)(z^2-4) + 4\kappa - 8}}{2}, \quad x' = \frac{yz + \sqrt{(y^2-4)(z^2-4) + 4\kappa - 8}}{2}$$

and therefore

$$x' - x = \sqrt{(y^2-4)(z^2-4) + 4\kappa - 8}.$$

Now from our inequalities above we obtain either $y_n z_n - x_n > 2$, giving us two coordinates greater than 2, or

$$(x_{n+1}, y_{n+1}, z_{n+1}) = (y_n, x'_n, z_n).$$

We see that z_n only ceases to be constant when some other coordinate becomes greater than 2, ending the algorithm. The sum of the other two coordinates $x_n + y_n$ is strictly increasing and

$$\begin{aligned} (x_{n+1} + y_{n+1}) - (x_n + y_n) &= x'_n - x_n \\ &= \sqrt{(y_n^2-4)(z_n^2-4) + 4\kappa - 8}. \end{aligned}$$

Now $0 < y_n \leq 2$ and $z > 2$, so that the product $(y_n^2 - 4)(z_n^2 - 4)$ is negative. The factor $(z_n^2 - 4)$ is a positive constant, and the other factor $y_n^2 - 4$ increases towards 0 as y_n increases. Thus the product $(y_n^2 - 4)(z_n^2 - 4)$ increases with n and

$$\begin{aligned} x_n' - x_n &= \sqrt{(y_n^2 - 4)(z_n^2 - 4) + 4\kappa - 8} \\ &\geq \sqrt{(y_0^2 - 4)(z_0^2 - 4) + 4\kappa - 8} \end{aligned}$$

which is positive as it is the discriminant of the quadratic with $x_0 \neq x_0'$ as roots. Thus $x_n + y_n$ increases by at least this amount each time. After a finite number of moves then $x_n + y_n > 4$, so at least one of x_n, y_n becomes larger than 2.

The region R_3

Here we may assume $-2 \leq x \leq 2 < y < z$, possibly after reordering. Now simply take $(x', y', z') = (yz - x, y, z)$. Clearly $y, z > 2$ and $x' = yz - x > 2 \times 2 - 2 = 2$. So $(x', y', z') \in R_1$.

The region R_2

Applying a sign change manoeuvre, we have $x, y, z < -2$. Now we apply a Markoff move and a sign change

$$(x, y, z) \mapsto (yz - x, y, z) \mapsto (yz - x, -y, -z).$$

Clearly $-y, -z > 2$ and $yz > 4$, $-x > 2$ imply $yz - x > 6 > 2$. Thus $(x', y', z') \in R_1$.

This concludes the proof of lemma 7.6.5 and hence proposition 7.6.2.

Note that in fact, the change of basis can be taken to be orientation-preserving. If necessary, we simply make the change of basis $(G, H) \mapsto (H, G)$ say, which on $X(S)$ maps $(x, y, z) \mapsto (y, x, z)$.

7.6.3 Explicit construction

We now have a basis G, H of $\pi_1(S)$ such that, with $\rho(G) = g, \rho(H) = h$, we have

$$(\operatorname{Tr} g, \operatorname{Tr} h, \operatorname{Tr} gh) \in (2, \infty)^3,$$

so that g, h, gh are hyperbolic isometries of $\mathbb{H}^2$. We will first explain the significance of the fact that all traces are positive.

From lemma 3.2.2 we see that the axes of g and h are disjoint. Since $\operatorname{Tr}[g, gh] = \operatorname{Tr}[h, gh] = \operatorname{Tr}[g, h]$, the axes of g, h, gh are all disjoint. These axes cannot share a fixed point at infinity either, for then $\operatorname{Tr}[g, h] = \pm 2$.

We will rely on results of Gilman and Maskit in [24]. Let $C(g,h)$ denote the cross ratio

$$C(g,h) = \frac{(r_g - a_h)(a_g - r_h)}{(r_g - r_h)(a_g - a_h)}$$

in the upper half plane model (here r_g and a_g respectively denote the repulsive and attractive fixed points of g, as in chapter 3). This quantity just tells us the orientation of the axes of g, h with respect to each other. (Note this is the reciprocal of the definition in [24]; but the definition in that paper conflicts with their theorem; and certainly with their figure 2. Rewriting their definition of cross-ratio seems better than rewriting their theorem.)

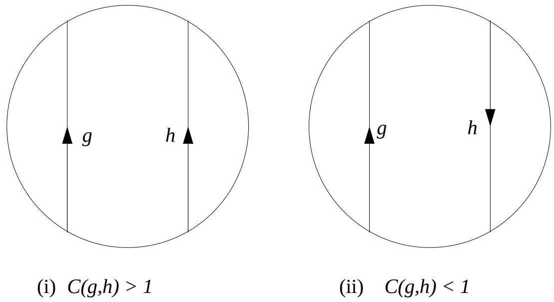

Figure 7.19: Different respective orientations of axes of g, h.

Lemma 7.6.7 *Suppose $g, h \in PSL_2\mathbb{R}$ are hyperbolic and $\mathrm{Tr}[g,h] > 2$. Then $C(g,h) \in (1,\infty)$ iff the axes of g, h are oriented as in figure 7.19(i), and $C(g,h) \in (0,1)$ iff the axes are oriented as in figure 7.19(ii).*

Proof In the situation of figure 7.19(i) we may project to the upper half plane as in figure 7.20. With lengths along the real axis α, β, γ as labelled then we have

$$C(g,h) = \frac{(r_g - a_h)(a_g - r_h)}{(r_g - r_h)(a_g - a_h)} = \frac{(\beta+\gamma)(\alpha+\beta)}{\beta(\alpha+\beta+\gamma)} > 1.$$

The inequality follows since $(\beta+\gamma)(\alpha+\beta) = \alpha\beta + \alpha\gamma + \beta^2 + \beta\gamma > \alpha\beta + \beta^2 + \beta\gamma = \beta(\alpha+\beta+\gamma)$. In the situation of figure 7.19(ii) a similar computation gives $C(g,h) \in (0,1)$. ∎

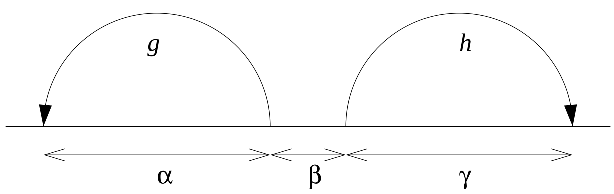

Figure 7.20: $C(g,h) > 1$, in the upper half plane.

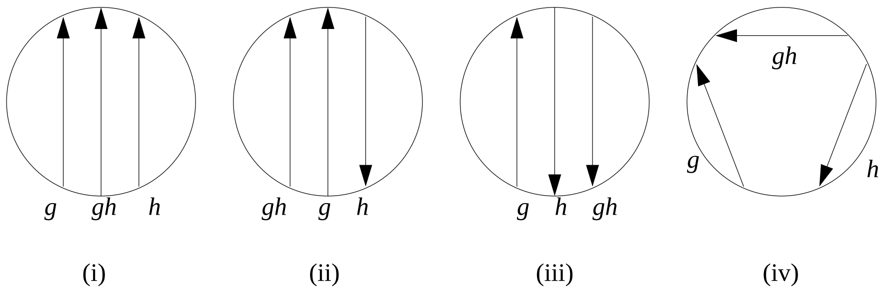

Figure 7.21: The possible arrangements of axes of g, h, gh.

Lemma 7.6.8 *Let* $g, h \in PSL_2\mathbb{R}$ *where* g, h, gh *are hyperbolic and* $\mathrm{Tr}[g,h] > 2$*. The possible arrangements of the axes of* g, h, gh *are shown in figure 7.21, and correspond to the following descriptions:*

(i) $C(g,h) \in (1,\infty)$ *and* $\mathrm{Tr}(g)\,\mathrm{Tr}(h)\,\mathrm{Tr}(gh) > 8$*;*

(ii) $C(g,h) \in (0,1)$ *and* $\mathrm{Tr}(g)\,\mathrm{Tr}(h)\,\mathrm{Tr}(gh) > 8$*;*

(iii) $C(g,h) \in (0,1)$ *and* $\mathrm{Tr}(g)\,\mathrm{Tr}(h)\,\mathrm{Tr}(gh) > 8$*;*

(iv) $C(g,h) \in (0,1)$ *and* $\mathrm{Tr}(g)\,\mathrm{Tr}(h)\,\mathrm{Tr}(gh) < -8$.

The proof will use the following theorem of Gilman and Maskit.

Theorem 7.6.9 (Gilman–Maskit [24]) *Let* g, h *be hyperbolic isometries such that* g, h *have no fixed points in common, the axes of* g *and* h *do not intersect, and* gh *is also hyperbolic.*

(i) If $C(g,h) \in (1,\infty)$ *then* $\mathrm{Tr}(g)\,\mathrm{Tr}(h)\,\mathrm{Tr}(gh) > 8$.

(ii) If $C(g,h) \in (0,1)$ *then* $\mathrm{Tr}(g)\,\mathrm{Tr}(h)\,\mathrm{Tr}(gh) < -8$ *if and only if the axes of* g, h, gh *bound a common region in* $\mathbb{H}^2$. ■

Proof (of lemma 7.6.8) We already know about the cross ratios. Suppose the axes of g, h are oriented as in figure 7.21(i), so that $C(g,h) > 1$ and by theorem 7.6.9, $\mathrm{Tr}(g)\,\mathrm{Tr}(h)\,\mathrm{Tr}(gh) > 8$. We use the result that a hyperbolic isometry translating a distance d along an axis l is the composition of two reflections, in lines perpendicular to l spaced $d/2$ apart. Denote by R_l the reflection in the line l. Let β denote the common perpendicular of $\mathrm{Axis}\, g$ and $\mathrm{Axis}\, h$, and then choose perpendiculars α, γ so that $g = R_\gamma R_\beta$ and $h = R_\beta R_\alpha$. Then $gh = R_\gamma R_\alpha$, and gh is hyperbolic. Thus α, γ do not intersect, and their common perpendicular is the axis of gh. As $\mathrm{Tr}[g, gh] = \mathrm{Tr}[h, gh] = \mathrm{Tr}[g,h] > 2$, by lemma 3.2.2, $\mathrm{Axis}\, gh$ is disjoint from $\mathrm{Axis}\, g$

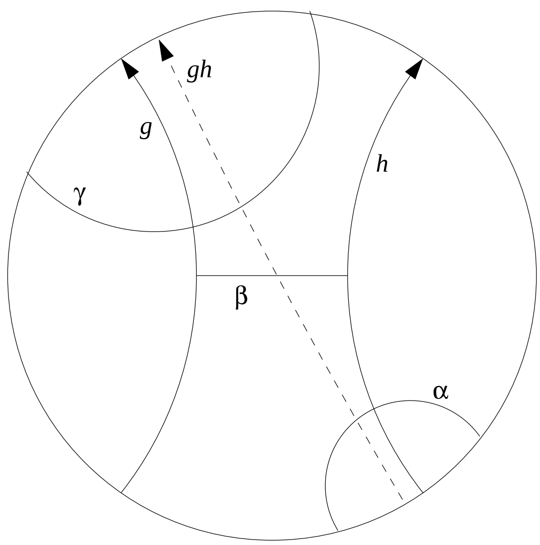

Figure 7.22: The situation when $C(g,h) > 1$.

and Axis h. We see in this situation, as shown in figure 7.22, that Axis gh must pass through the region bounded by Axis g and Axis h, with the orientation shown. It follows that gh must lie as shown in figure 7.21(i).

Now suppose that the axes of g, h are oriented with $C(g,h) < 1$, as in figures 7.21(ii)–(iv). Let α, β, γ be perpendiculars as before. We see by varying the possible positions of α and γ, and noting that Axis gh must be disjoint from Axis g and Axis h, that there are precisely three possible locations for Axis gh, namely those shown. By theorem 7.6.9, $\operatorname{Tr}(g)\operatorname{Tr}(h)\operatorname{Tr}(gh) < -8$ in case (iv) and > 8 in cases (ii) and (iii). ∎

Returning to the problem at hand, we have a basis with $\operatorname{Tr} g, \operatorname{Tr} h, \operatorname{Tr} gh > 2$. So lemma 7.6.8 tells us that the cases we must consider are precisely those in figure 7.21(i),(ii),(iii). We will explicitly show how to choose p so that $\mathcal{P}(g,h,p)$ is a non-degenerate simple pentagon bounding an embedded disc.

Case (i)

Assume g, h, gh have axes as shown in figure 7.21(i). Note that $hg = h(gh)h^{-1} = g^{-1}(gh)g$ so the axis of hg is the image of the axis of gh under either of h or g^{-1}. Thus r_{hg} must lie between r_{gh} and r_g; and a_{hg} must lie between a_{gh} and a_h. So Axis hg is arranged as shown in figure 7.23.

Let r be the intersection of the axes of gh and hg, and let $p = h^{-1}g^{-1}(r)$. Then

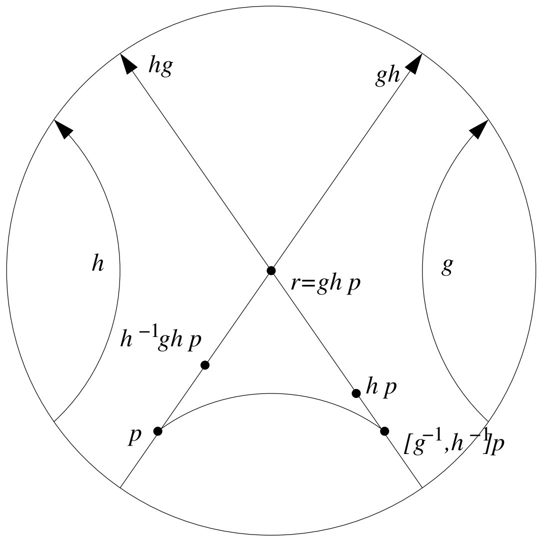

Figure 7.23: Construction in case (i).

we have immediately

$$\begin{aligned} p &\in \operatorname{Axis} gh = \operatorname{Axis} h^{-1}g^{-1} \\ hp &\in h(\operatorname{Axis} h^{-1}g^{-1}) = \operatorname{Axis} g^{-1}h^{-1} = \operatorname{Axis} hg \\ ghp &= r = \operatorname{Axis}(hg) \cap \operatorname{Axis}(gh) \\ h^{-1}ghp &\in h^{-1}(\operatorname{Axis} hg) = \operatorname{Axis} gh \\ [g^{-1},h^{-1}]p &\in g^{-1}(\operatorname{Axis} gh) = \operatorname{Axis} hg. \end{aligned}$$

Since $p = (gh)^{-1}(r)$, p lies on Axis gh on the same side of r as r_{gh}. Similarly $[g^{-1},h^{-1}]p = (hg)^{-1}(r)$ lies on Axis hg on the same side of r as r_{hg}. Considering the action of h, we see that $h^{-1}ghp = h^{-1}(r)$ lies on Axis gh on the same side of r as r_{gh}; and similarly hp lies on Axis hg on the same side of r as r_{hg}. Further, since h maps the directed segment $(h^{-1}ghp, p)$ to the directed segment (ghp, hp), we see that $h^{-1}ghp$ lies on the same side of p as r. Similarly, as g maps the directed segment $([g^{-1},h^{-1}]p, hp)$ to $(h^{-1}ghp, ghp)$, we see hp lies on the same side of $[g^{-1},h^{-1}]p$ as r. So $\mathcal{P}(g,h,p)$ indeed appears as in figure 7.23, and it is a non-degenerate pentagon bounding an embedded disc.

Examining figure 7.23, we see that $\mathcal{P}(g,h,p)$ contains two straight angles, so $\theta \in (2\pi, 3\pi)$. From corollary 3.7.3 $\operatorname{Tr}[g,h] > 2$ implies $[g,h] \in \operatorname{Hyp}_0$, so we have by proposition 3.4.4 $\operatorname{Tw}([g^{-1},h^{-1}],p) \in (-\pi,\pi)$. From lemma 5.4.1 we have $\theta \equiv \pi \pm \operatorname{Tw}([g^{-1},h^{-1}],p)$ modulo 2π, respectively according to the orientation of S. Thus $\theta = 3\pi \pm \operatorname{Tw}([g,h],p)$, with the $+$ or $-$ taken according to the orientation of S. It follows from the conditions of lemma 5.4.1 that $\operatorname{Tw}([g^{-1},h^{-1}],p) > 0$ if and only if

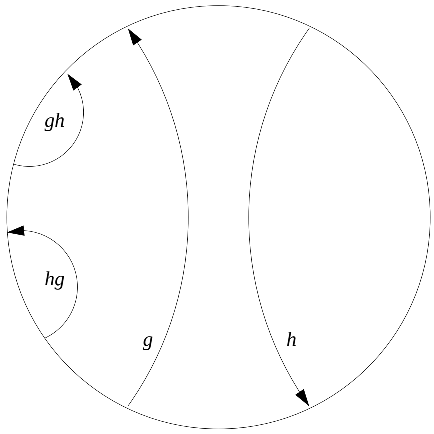

Figure 7.24: Axes of g, h, gh, hg in case (ii).

$p \to [g^{-1}, h^{-1}]p$ bounds $\mathcal{P}(g,h,p)$ on its left, i.e. ∂S traversed in the direction of $[G,H]$ bounds S on its left.

Case (ii)

Now take g, h, gh arranged as in figure 7.21(ii). Again we consider the axis of hg, which is the image of $\text{Axis}\, gh$ under h or g^{-1}, and therefore must lie on the same side of $\text{Axis}\, g$ as $\text{Axis}\, gh$. The axes of gh and hg may or may not intersect; we do not care. See figure 7.24.

In particular $h^{-1}(a_{hg}) = a_{gh}$. Thus $h^{-1}(r_g)$ lies between r_g and a_{gh}, in the same arc of the circle at infinity as a_{hg}. Also $h^{-1}(a_g)$ lies in the arc between a_g and r_h. Since $h^{-1}(\text{Axis}\, g) = \text{Axis}\, h^{-1}gh$, we have the arrangement of axes as shown in figure 7.25

Let $r = \text{Axis}\, g \cap \text{Axis}\, h^{-1}gh$, and let $p = h^{-1}r$. Then we have immediately:

$$\begin{aligned} p &\in h^{-1}(\text{Axis}\, g) = \text{Axis}\, h^{-1}gh \\ hp = q &= \text{Axis}\, g \cap \text{Axis}\, h^{-1}gh \\ ghp &\in \text{Axis}\, g \\ h^{-1}ghp &\in h^{-1}(\text{Axis}\, g) = \text{Axis}\, h^{-1}gh \end{aligned}$$

Now considering the action of h, we see that $p = h^{-1}r$ lies to the same side of r as $h^{-1}ghp$. Since h maps the directed segment $(h^{-1}ghp, p)$ to (ghp, hp), we see that $h^{-1}ghp$ lies on the opposite side of p as $hp = r$. Now $h^{-1}ghp$ lies to the right of $\text{Axis}\, g$ in the diagram shown, so $[g^{-1}, h^{-1}]p$ lies to the right of $\text{Axis}\, g$ also. And considering the action of g^{-1}, we see that the image of the axis of $h^{-1}gh$ under g^{-1} is disjoint from $\text{Axis}\, h^{-1}gh$ and lies below it. So $[g^{-1}, h^{-1}]p$ lies to the right of

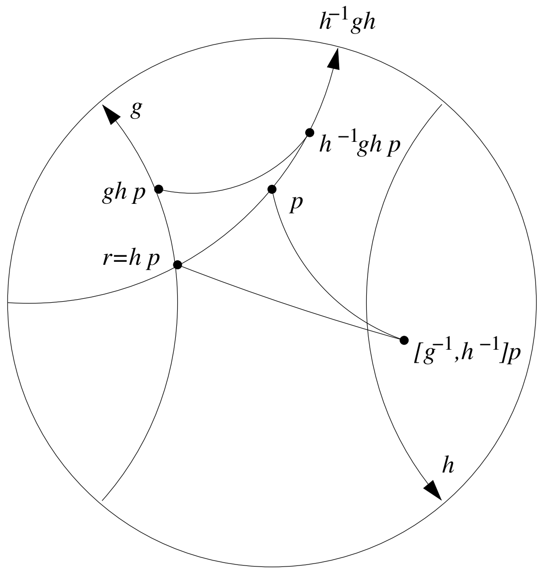

Figure 7.25: Construction in case (ii).

Axis g and also below Axis $h^{-1}gh$. Hence $\mathcal{P}(g,h,p)$ is as shown in figure 7.25, and is non-degenerate, bounding an embedded disc.

Examining figure 7.25, we may chase around unit vectors and see that $\mathrm{Tw}([g^{-1},h^{-1}],p) < 0$. By lemma 5.4.1 we have $\theta \equiv \pi + \mathrm{Tw}([g^{-1},h^{-1}],p) \bmod 2\pi$. As $\mathcal{P}(g,h,p)$ contains a reflex angle, $\theta \in (\pi, 3\pi)$. So we must have $\theta = 3\pi + \mathrm{Tw}([g^{-1},h^{-1}],p)$ in this case. If the opposite orientation occurs then $\mathrm{Tw}([g^{-1},h^{-1}],p) > 0$ and $\theta = 3\pi - \mathrm{Tw}([g^{-1},h^{-1}],p)$.

Case (iii)

The final case is not very different to case (ii). Again we consider the location of Axis hg. By a similar argument as in case (ii), we deduce that Axis hg lies on the same side of Axis h as Axis gh. Thus, similarly to case (ii), we deduce that Axis $g^{-1}hg$ lies as shown in figure 7.26. Let r be the intersection of Axis h and Axis $g^{-1}hg$, and let $p = h^{-1}g^{-1}hr$. Then we have $h^{-1}ghp = r$, so $ghp \in$ Axis h in the direction shown in figure 7.26. The segment $hp \to [g^{-1},h^{-1}]p$ is the image of $ghp \to h^{-1}ghp$ under g^{-1}, hence is a segment on g^{-1} Axis $h =$ Axis $g^{-1}hg$ in the arrangement shown in figure 7.26. Finally as hp lies to the left of Axis h, p lies to the left of Axis h, translated along the constant distance curve from Axis h through hp. It follows that p lies above Axis $g^{-1}hg$. So $\mathcal{P}(g,h,p)$ lies as shown and is non-degenerate, bounding an embedded disc.

By lemma 5.1.5, we conclude in each case that ρ is the holonomy of a hyperbolic cone-manifold structure on S with no interior cone points and at most one corner

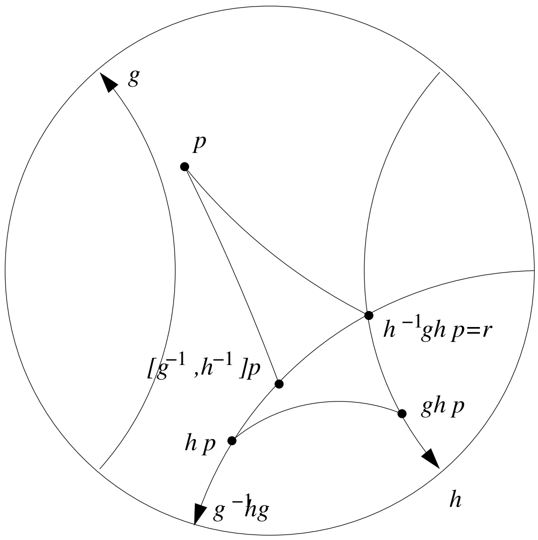

Figure 7.26: Construction in case (iii).

point. By the same argument as in the previous case, $\theta \in (\pi, 3\pi)$ and $\theta = 3\pi \pm \mathrm{Tw}([g^{-1}, h^{-1}], p)$, according to the orientation of S.

This completes the proof of proposition 7.6.3, and indeed of theorem A.

Chapter 8

Higher genus surfaces

8.1 Statements and discussion

We now turn to higher genus surfaces. First we will summarise the theorems and their proofs.

In section 8.2 we will prove the following theorem [25]:

Goldman's Theorem *Let S be an orientable surface with $\chi(S) < 0$, and let ρ be a representation $\pi_1(S) \longrightarrow PSL_2\mathbb{R}$. If S has boundary, assume ρ takes each boundary curve to a non-elliptic element, so the relative Euler class $\mathcal{E}(\rho)$ is well-defined. A representation ρ is the holonomy of a complete hyperbolic structure on S with totally geodesic or cusped boundary components (respectively as the boundary curve is hyperbolic or parabolic) if and only if $\mathcal{E}(\rho)[S] = \pm\chi(S)$.*

Recall, as discussed in section 4.2, why this statement makes sense. For a chosen basepoint $p \in s$ and boundary curve C, any two choices of loops based at p homotopic to C are conjugate in the fundamental group, hence we can reasonably say that $\rho(C)$ is non-elliptic. We will see along the way that $\mathcal{E}(\rho)[S] = \pm\chi(S)$ implies that no boundary curve can be taken by ρ to the identity, so $\rho(C)$ non-elliptic means that $\rho(C)$ is hyperbolic or parabolic.

One direction is clear, by proposition 4.2.2: if ρ is the holonomy representation of a complete hyperbolic structure of the type discussed then $\mathcal{E}(\rho)[S] = \pm\chi(S)$. In section 8.2 we will take a representation with $\mathcal{E}(\rho)[S] = \pm\chi(S)$ and construct a hyperbolic structure on S. In the special case of a closed surface this becomes:

Corollary *Let S be a closed orientable surface of genus $g \geq 2$. A representation $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ is the holonomy of a complete hyperbolic structure on S if and only if $\mathcal{E}(\rho)[S] = \pm\chi(S)$* ■

We are masters of punctured tori; we will need to become masters of our pants as well, in order to proceed. For we will take a decomposition of S into pants and punctured tori, and show that each part has a complete hyperbolic structure induced by ρ. In section 8.2.1 we will show how to decompose our surface; in section 8.2.2 we will obtain hyperbolic structures on our pants; in section 8.2.3 we will put the pieces together geometrically.

In the case of a genus 2 surface, we can prove the following result for representations with $\mathcal{E}(\rho)[S] = \pm 1$.

Theorem B *Let S be a genus 2 closed surface. Let $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ be a representation with $\mathcal{E}(\rho)[S] = \pm 1$. Suppose that there is a separating curve C on S such that $\rho(C)$ is not hyperbolic. Then ρ is the holonomy of a hyperbolic cone-manifold structure on S with one cone point of angle 4π.*

We will prove this result in section 8.3: we can split a genus 2 surface into two punctured tori, and work with both pieces. If one could prove that for all representations with $\mathcal{E}(\rho)[S] = \pm 1$, there is a separating curve C which is not hyperbolic, then we would clearly have a stronger result. However I have not been able to do so, using the low-level techniques at hand.

We will also prove the following result for a general closed surface of genus g:

Theorem C *Let S be a closed orientable surface of genus $g \geq 2$. Let $Y \subset X(S)$ denote the set of characters of representations $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ such that*

(i) $\mathcal{E}(\rho)[S] = \pm(\chi(S) + 1)$;

(ii) there exists a non-separating simple closed curve C on S_1 such that $\rho(C)$ is elliptic.

Almost every ρ with character in Y is the holonomy of a cone-manifold structure on S with a single cone point with cone angle 4π.

Here the measure on $X(S)$ is μ_S as described in section 4.4. By the theorem of Goldman 4.4.1, the sets of representations with $\mathcal{E}(\rho)[S] = \chi(S) + 1$ or $-\chi(S) + 1$ form connected components of $R(S)$; these characters form connected components of $X(S)$. An abelian representation has $\mathcal{E}(\rho)[S] = 0$, by lemma 4.3.3, so there are no singular points to consider in the situation of the theorem. The "almost every" condition arises since the proof uses ergodicity properties of the action of the mapping class group on the character variety described in sections 4.5 and 6.3.

Our proof relies on the existence of a non-separating simple closed curve with elliptic holonomy; this can then be cut off, allowing us to localise the deficiency in the Euler class. But we have not been able to show such a curve exists in general.

Question 8.1.1 *For a general closed surface of genus $g \geq 2$, is almost every representation with Euler class $\pm(\chi(S)+1)$ a holonomy representation? What about surfaces with boundary, and relative character varieties?*

Nor is it clear whether the word "almost" could be removed from the statement of the theorem. According to Tan [58], Goldman and Neumann had an unpublished proof that for any closed surface S of genus ≥ 2, and a representation ρ with $\mathcal{E}(\rho)[S] = \pm(\chi(S)+1)$, ρ is the holonomy of a hyperbolic cone-manifold structure with a single cone point with angle 4π. However this result now appears to be in doubt. We know of no counterexample; the question remains open.

Question 8.1.2 *Is every representation of a general genus $g \geq 2$ closed surface with Euler class $\pm(\chi(S)+1)$ a holonomy representation? What about surfaces with boundary?*

Our ideas rely heavily on the assumption that $\mathcal{E}(\rho)$ is close to extremal; this guarantees us that once we localise the deficiency in $\mathcal{E}(\rho)$, we can cut it off, and the representation on the other side will have extremal Euler class. For other values of $\mathcal{E}(\rho)$, the question remains how prevalent the holonomy representations are. There are clearly none for $\mathcal{E}(\rho)[S] = 0$: by proposition 4.2.2 this implies the orders of cone points satisfy $\sum s_i = -\chi(S)$, which contradicts Gauss-Bonnet 2.2.2. In [58], Tan gives a construction of a representation of a genus 3 closed surface S (so $\chi(S) = -4$), with Euler class $\mathcal{E}(\rho)[S] = 2$, which is not the holonomy of any cone-manifold structure. This representation simply pinches down a handle, mapping it to the identity. But Tan also finds representations arbitrarily close to this one, which do give branched hyperbolic structures. So we can still ask:

Question 8.1.3 *For a given integer $m \neq 0$, $\chi(S)+1 \leq m \leq -\chi(S)-1$, are holonomy representations dense, or conull, in the set of representations with fixed Euler class $\mathcal{E}(\rho)[S] = m$?*

8.2 Goldman's theorem

In this section let S be an orientable surface of genus g, with n boundary components. Assume ρ is non-elliptic on each boundary component, so $\mathcal{E}(\rho)$ is well-defined. Assume $\mathcal{E}(\rho)[S] = -\chi(S) > 0$; the case $\mathcal{E}(\rho)[S] = \chi(S) < 0$ is similar with reversed orientation.

8.2.1 Splitting up is hard to do

In [23], Gallo, Kapovich and Marden show that for any non-elementary representation $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{C}$, where S is a closed oriented surface with $\chi(S) < 0$, there exists a system of disjoint curves C_i decomposing S into pants P_j, such that the restriction of ρ to each P_j is a 2-generator classical Schottky group. The proof is long, but their methods apply immediately to the case of representations $\pi_1(S) \longrightarrow PSL_2\mathbb{R}$, and when there are boundary components. That the restriction of ρ to each P_j is a 2-generator classical Schottky group implies that each $\rho(C_i)$ is hyperbolic. An elementary representation can be represented by diagonal matrices, hence lies in the same topological component of the representation space as the identity, hence has Euler class zero by theorem 4.4.1.

The proof of the theorem relies upon applying Dehn twists to obtain sufficiently "complicated" curves that they have holonomy with large trace. Algorithmically, it cuts $g - 1$ "handles", one at a time, so that the genus decreases by 1 at each stage; and from the remaining piece of genus 1, cuts off pants (choosing pairs of boundary circles to form into pants arbitrarily each time) until the genus 1 piece is just a once-punctured torus; then this too is cut into pants. But since we are so comfortable with punctured tori, we could perform the algorithm so g of the pants have pairs of boundary curves identified, and we glue them back together to give punctured tori. So we can decompose S along curves with hyperbolic holonomy into g tori and $g + n - 2$ pants.

Then we can assume the surfaces combinatorially fit together as in figure 8.1. If S is closed of genus 2, then we just have two punctured tori. Otherwise, none of the punctured tori are adjacent. We draw all the punctured tori leftmost; these must then be connected together. If S has no boundary, we simply connect up all the punctured tori by pants. If S has boundary, we may add on further pants to the situation of 8.1 to obtain more boundary components.

In short: their theorem trivially implies the following.

Theorem 8.2.1 (Gallo, Kapovich, Marden [23]) *Let S be an oriented surface with $\chi(S) < 0$ and let $\rho : \pi_1(S) \longrightarrow PSL_2\mathbb{R}$ be a representation with $\mathcal{E}(\rho)[S]$ well-defined and equal to $\pm\chi(S)$. Then there exists a system of disjoint curves C_i decomposing S into pants and punctured tori, such that each $\rho(C_i)$ is hyperbolic.* ■

Consider the fundamental groups of S, and the subsurfaces S_i into which it is decomposed. We need to specify basepoints q_i on each S_i, and a basepoint q on S. On each punctured torus we specify a basepoint on the boundary, as in chapter 7. On each pair of pants we arbitrarily specify a basepoint. We arbitrarily choose

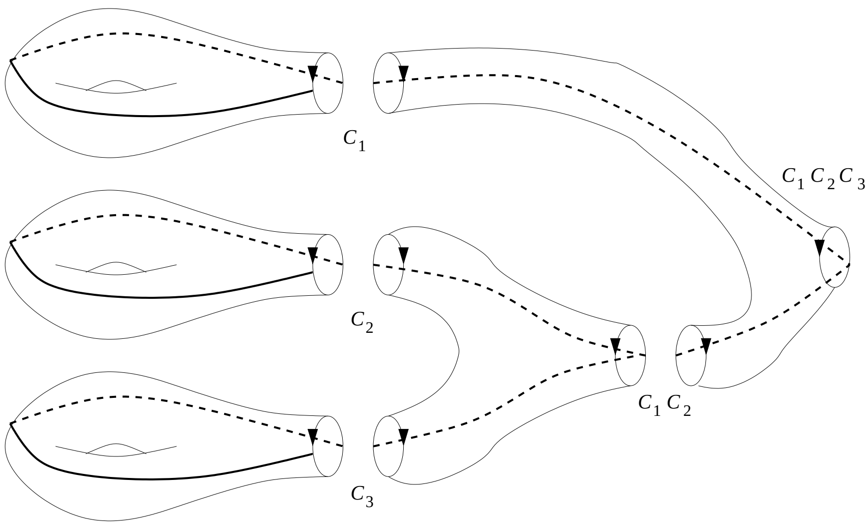

Figure 8.1: Connecting up surfaces, and fundamental group.

one of the q_i to be the basepoint q for S. To specify how the fundamental groups $\pi_1(S_i, q_i)$ and $\pi_1(S, q)$ relate to each other, we must choose paths between endpoints. We take a combinatorial tree T dual to the decomposition of S, with one vertex for each q_i. Then we choose a map $\mathcal{T} : T \longrightarrow S$ such that $\mathcal{T}$ maps the vertices of T onto the corresponding q_i. This gives well-defined paths between the q_i. We have inclusions $\iota_i : \pi_1(S_i, q_i) \hookrightarrow \pi_1(S, q_i)$ (note basepoints here). Let α_i be the unique path from q to each q_i along the tree $\mathcal{T}$, then we have isomorphisms

$$\zeta_i : \pi_1(S, q_i) \xrightarrow{\cong} \pi_1(S, q), \quad x \mapsto \alpha_i.x.\alpha_i^{-1}.$$

Suppose ρ is the holonomy of some geometric structure on S with basepoint q, corresponding to a chosen lift $\tilde{q} \in \tilde{S}$. Then there is a preferred lift $\tilde{\mathcal{T}}$ and hence a preferred lift $\tilde{q}_i$ of each q_i. We see that $\rho \circ \zeta_i$ will be the holonomy of an isometric geometric structure on S with basepoint q_i, corresponding to the lift $\tilde{q}_i$. Thus for each S_i we have a representation $\rho_i : \pi_1(S_i, q_i) \longrightarrow PSL_2\mathbb{R}$ given by the composition

$$\pi_1(S_i, q_i) \xrightarrow{\iota_i} \pi_1(S, q_i) \xrightarrow{\zeta_i} \pi_1(S, q) \xrightarrow{\rho} PSL_2\mathbb{R}.$$

Returning to the global representation ρ, each $\rho(C_i)$ is hyperbolic, and each boundary curve is non-elliptic. So we have well-defined relative Euler classes $\mathcal{E}(\rho_i)$, and they are additive by lemma 4.2.1:

$$\sum_{i=1}^{-\chi(S)} \mathcal{E}(\rho_i)[S_i] = \mathcal{E}(\rho)[S] = -\chi(S).$$

Since by the Milnor-Wood inequality 4.3.2 $|E(\rho_i)[S_i]| \leq 1$, we must have $\mathcal{E}(\rho_i)[S_i] = 1$ for each i.

Consider an S_i which is a punctured torus. If $S_i = S$ then we are done by the previous chapter. Otherwise ∂S_i is some decomposition curve. Taking any basis G_i, H_i for $\pi_1(S_i, q_i)$ we have $[\rho_i(G_i), \rho_i(H_i)]$ hyperbolic, and by proposition 4.3.4 $\text{Tr}[\rho_i(G_i), \rho_i(H_i)] < -2$. By section 7.2 ρ_i is the holonomy of a complete hyperbolic structure on S_i with totally geodesic boundary. We must consider those S_i which are pairs of pants.

8.2.2 How to hyperbolize your pants

Let S_i denote a pair of pants, and let C_1, C_2, C_3 denote elements of $\pi_1(S_i)$ homotopic to the boundary curves so that $C_1C_2C_3 = 1$. Thus $\pi_1(S_i) = \langle C_1, C_2, C_3 \mid C_1C_2C_3 = 1\rangle$. Let $\rho_i(C_j) = c_j$, so each c_j is non-elliptic; the C_j which correspond to decomposition curves are hyperbolic. Recall $\mathcal{E}(\rho_i)[S_i] = 1$. By lemma 4.3.3, this implies ρ_i is non-abelian. Hence no c_i is the identity in $PSL_2\mathbb{R}$: each c_j is hyperbolic or parabolic. Thus each c_j has a simplest lift $\tilde{c}_j \in \text{Hyp}_0 \cup \text{Par}_0$. Since $\mathcal{E}(\rho_i)[S_i] = 1$ we have by 4.3.1, $\tilde{c}_1\tilde{c}_2\tilde{c}_3 = \mathbf{z}$, so $\Theta(\tilde{c}_1\tilde{c}_2\tilde{c}_3) = \pi$.

Lemma 8.2.2 $\Theta(\tilde{c}_1\tilde{c}_2) + \Theta(\tilde{c}_3) = \pi$.

Proof Each $\tilde{c}_j \in \text{Hyp}_0 \cup \text{Par}_0$ so $\Theta(\tilde{c}_j) \in (-\pi/2, \pi/2)$, by 3.6.4. By approximate additivity of Θ (3.5.1) we have

$$|\Theta(\tilde{c}_1\tilde{c}_2) - \Theta(\tilde{c}_1) - \Theta(\tilde{c}_2)| < \frac{\pi}{2}$$

so $|\Theta(\tilde{c}_1\tilde{c}_2)| < 3\pi/2$. Now $c_1c_2 = c_3^{-1}$ so $\tilde{c}_1\tilde{c}_2 = \mathbf{z}^m\tilde{c}_3^{-1}$ for some integer m and $\Theta(\tilde{c}_1\tilde{c}_2) = m\pi - \Theta(\tilde{c}_3)$. Since $\Theta(\tilde{c}_1\tilde{c}_2) \in (-3\pi/2, 3\pi/2)$ and $\Theta(\tilde{c}_3) \in (-\pi/2, \pi/2)$ we have $m\pi = \Theta(\tilde{c}_1\tilde{c}_2) + \Theta(\tilde{c}_3) \in (-2\pi, 2\pi)$, so $m = -1, 0, 1$. Again by approximate additivity

$$|\Theta(\tilde{c}_1\tilde{c}_2\tilde{c}_3) - \Theta(\tilde{c}_1\tilde{c}_2) - \Theta(\tilde{c}_3)| = |\pi - m\pi| < \frac{\pi}{2},$$

so we have $m = 1$, as desired. ■

Lemma 8.2.3 *The product* $\tilde{c}_1\tilde{c}_2 \in \text{Hyp}_1 \cup \text{Par}_1$. *The product* $\text{Tr}(c_1)\,\text{Tr}(c_2)\,\text{Tr}(c_1c_2)$ *is well-defined and* ≤ -8.

Proof Note that lifting $c_j \in PSL_2\mathbb{R}$ to a matrix in $SL_2\mathbb{R}$ of either sign does not change the product of traces, so it is well-defined. We have $\tilde{c}_1, \tilde{c}_2\tilde{c}_3 \in \text{Hyp}_0 \cup \text{Par}_0$ so by 3.7.1, $\text{Tr}(\tilde{c}_1), \text{Tr}(\tilde{c}_2)\,\text{Tr}(\tilde{c}_3) \geq 2$ and by 3.6.4 $\Theta(\tilde{c}_1), \Theta(\tilde{c}_2), \Theta(\tilde{c}_3) \in (-\pi/2, \pi/2)$. From the previous lemma we have $\Theta(\tilde{c}_1\tilde{c}_2) = \pi - \Theta(\tilde{c}_3) \in (\pi/2, 3\pi/2)$. Since $c_1c_2 = c_3^{-1}$, which is hyperbolic or parabolic, by 3.6.4 again $\tilde{c}_1\tilde{c}_2 \in \text{Hyp}_1 \cup \text{Par}_1$. Then by 3.7.1 again $\text{Tr}(\tilde{c}_1\tilde{c}_2) \leq -2$. Thus the product of the traces is ≤ -8. ■

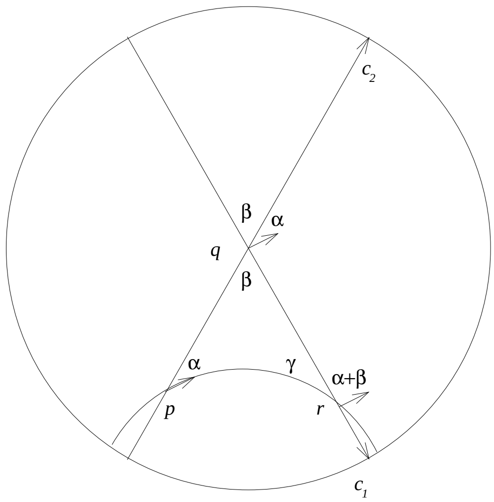

Figure 8.2: Unit vector chase if $\text{Tr}[c_1, c_2] < 2$.

Note that $\pi_1(S_i)$ also has presentation $\langle C_1, C_2 \rangle$, so ρ can be considered as a representation of the free group on two generators; in particular, our characterisation of reducible representations (proposition 6.1.3) and the results of section 6.5 apply.

Lemma 8.2.4 $\text{Tr}[c_1, c_2] > 2$.

PROOF Since by the previous lemma $\tilde{c}_1\tilde{c}_2 \in \text{Hyp}_1 \cup \text{Par}_1$, from 3.4.4 we have for any $p \in \mathbb{H}^2$, $\text{Tw}(\tilde{c}_1\tilde{c}_2, p) \in (\pi, 3\pi)$.

Suppose $\text{Tr}[c_1, c_2] < 2$, so that by lemma 3.2.2 c_1, c_2 are both hyperbolic and their axes intersect at a point $q \in \mathbb{H}^2$. Let $p = c_2^{-1}(q)$ and let $r = c_1(q)$. Let the angles in triangle pqr be α, β, γ as shown in figure 8.2, so $\alpha+\beta+\gamma < \pi$. We perform a unit vector chase commencing with the vector at p pointing towards r. Under the actions of $\tilde{c}_2$ and $\tilde{c}_1$, we obtain the vectors shown, so (taking into account the two possible orientations) $\text{Tw}(\tilde{c}_1\tilde{c}_2, p) = \pm(\pi - \alpha - \beta - \gamma) \in (-\pi, \pi)$, a contradiction.

If $\text{Tr}[c_1, c_2] = 2$ then by proposition 6.1.3 ρ_i is reducible. As ρ_i is non-abelian, lemma 6.5.2 describes c_1, c_2. Either one of c_1, c_2 is hyperbolic and the other parabolic, with a common fixed point; or both c_1, c_2 are hyperbolic, with exactly one shared fixed point. In both these cases, a similar unit vector chase contradicts $\tilde{c}_1\tilde{c}_2 \in \text{Hyp}_1 \cup \text{Par}_1$. ∎

For the next lemma, we consider the arrangement of the axes of the c_j. For hyperbolic c_j, this is well-defined; for parabolic c_j, consider the "axis" of c_j to degenerate to a point, namely the fixed point at infinity.

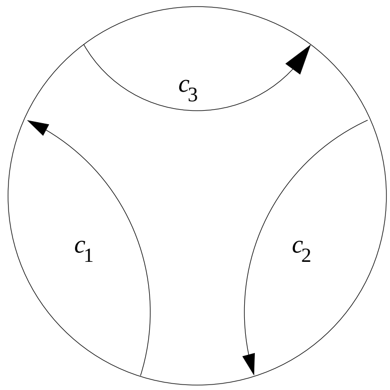

Figure 8.3: Arrangement of axes for an Euler-class-1 pants representation.

Lemma 8.2.5 *The axes of c_1, c_2, c_3, understood in the extended sense defined above, are disjoint and bound a common region, as in figure 8.3.*

PROOF First suppose all c_j are hyperbolic. Then we have $\mathrm{Tr}[c_1, c_2] > 2$ and also $\mathrm{Tr}(c_1)\,\mathrm{Tr}(c_2)\,\mathrm{Tr}(c_1c_2) < -8$, so we may apply lemma 7.6.8 and conclude that the axes are disjoint and bound a common region. Performing a unit vector chase shows that figure 8.3 depicts the situation $\tilde{c}_1\tilde{c}_2\tilde{c}_3 = \mathbf{z}$, and not its mirror reverse.

If some c_j are parabolic, then ρ is the limit of representations where all c_j are hyperbolic. So by continuity we obtain the desired result. ∎

Now we can construct a complete hyperbolic structure on our pants S_i. If c_1 is hyperbolic then it is the composition of two reflections in lines perpendicular to $\mathrm{Axis}\, c_1$. If c_1 is parabolic then it is the composition of two reflections in lines through its fixed point at infinity. The same applies to c_2. We may take one of these lines to be the common perpendicular of $\mathrm{Axis}\, c_1$ and $\mathrm{Axis}\, c_2$, or if c_j is parabolic then we take this line to run to the fixed point at infinity of c_j. Then $c_1c_2 = c_3^{-1}$ is the composition of two reflections, and hence $\mathrm{Axis}\, c_1c_2 = \mathrm{Axis}\, c_3$, or $\mathrm{Fix}\, c_3$, is as shown in figure 8.4.

Note that the (possibly degenerate) octagon shown in the figure has two pairs of sides identified under c_1, c_2, and $\pi_1(S_i) = \langle C_1, C_2 \rangle$, so the octagon forms a fundamental domain for a pair of pants. Since all the angles in the octagon are right angles or 0, the boundary edges wrap up to give geodesic boundary, or cusps. Considering the universal cover $\tilde{S}_i$ and a preferred lift $\tilde{q}_i$ of our basepoint, we obtain a partial developing map taking $\tilde{q}_i$ to some point in the fundamental domain. This developing map extends equivariantly to give a complete hyperbolic structure on S_i with totally geodesic or cusped boundary, accordingly as each c_j is hyperbolic or parabolic, with holonomy ρ. In fact the representation ρ is discrete and S is given by a quotient of the convex core of ρ.

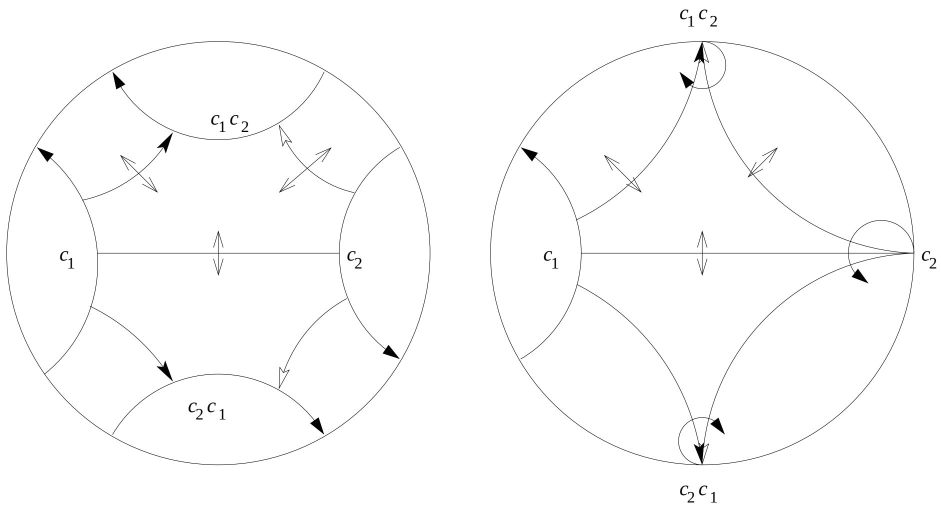

Figure 8.4: Fabricating pants.

We can make the following remark on orientation, recalling that $\mathcal{E}(\rho)[S_i] = 1$: if C_1, C_2, C_3 are the boundary curves with $C_1C_2C_3 = 1$, then in our developing map, the directed edges corresponding to C_j, i.e. the directed axis of c_j, bounds the fundamental domain on its right. If S_i inherits an orientation from $\mathbb{H}^2$, then the boundary curve C_j bounds S_i on its right also.

Proposition 8.2.6 *Let S_i denote a pair of pants with fundamental group having presentation $\pi_1(S_i) = \langle C_1, C_2, C_3 \mid C_1C_2C_3 = 1\rangle$. Let $\rho : \pi_1(S_i) \longrightarrow PSL_2\mathbb{R}$ be a representation taking each boundary curve C_i to a non-elliptic element, so the relative Euler class is well-defined. Suppose $\mathcal{E}(\rho)[S_i] = 1$ (resp. -1). Then ρ is the holonomy of a complete hyperbolic structure on S_i. Each C_i has hyperbolic or parabolic holonomy, and accordingly C_i is totally geodesic or cusped. Each C_i bounds S_i to its right (resp. left).* ∎

8.2.3 Putting the pieces together

We know how to construct a complete hyperbolic structure, with totally geodesic or cusped boundary, on each piece S_i. It only remains to see that these pieces fit together and that the holonomy on the surface S is ρ. We construct the hyperbolic structure on S piece by piece, starting from a first piece S_1 whose basepoint coincides with that of S, $q = q_1$. We then work outwards along the dual tree $\mathcal{T}$ to the decomposition. By our choice of decomposition (see figure 8.1), when adding a new piece, we only need to ensure it attaches along one boundary curve.

We will first consider the case where S has genus 0, i.e. S is an n-holed sphere. This decomposes into $n - 2$ pants. The combinatorial arrangement must be as in figure 8.1, minus the punctured tori. We may take a dual tree $\mathcal{T}$ as in that figure, and then we see that $\pi_1(S)$ has presentation

$$\langle C_1, \ldots, C_{n-2} \mid C_1 \cdots C_n = 1 \rangle.$$

Let $\rho(C_i) = c_i$. Since all the c_i are non-elliptic, they have preferred lifts $\tilde{c}_i$ to $\widetilde{PSL_2\mathbb{R}}$ and as $\mathcal{E}(\rho)[S] = -\chi(S) = n - 2$, by proposition 4.3.1 we have $\tilde{c}_1 \cdots \tilde{c}_n = \mathbf{z}^{n-2}$. Consider the following algebraic decomposition of the relator, corresponding to the decomposition of the surface.

$$\begin{aligned} 1 &= C_1 C_2 C_3 \cdots C_n \\ &= \left[C_1 C_2 (C_1 C_2)^{-1}\right] (C_1 C_2) C_3 C_4 \cdots C_n \\ &= \left[C_1 C_2 (C_1 C_2)^{-1}\right] \left[(C_1 C_2) C_3 (C_1 C_2 C_3)^{-1}\right] (C_1 C_2 C_3) C_4 \cdots C_n \\ &= \cdots \\ &= \left[C_1 C_2 (C_1 C_2)^{-1}\right] \left[(C_1 C_2) C_3 (C_1 C_2 C_3)^{-1}\right] \cdots \\ &\quad \left[(C_1 C_2 \cdots C_{n-3}) C_{n-2} (C_1 \cdots C_{n-2})^{-1}\right] \left[(C_1 C_2 \cdots C_{n-2}) C_{n-1} C_n\right] \end{aligned}$$

Each expression in square brackets is the relator in the presentation of the fundamental group of one of the pants in the decomposition in S. Now as each $\mathcal{E}(\rho_i)[S_i] = 1$, we have that relator equals $\mathbf{z}$. From proposition 8.2.6, each ρ_i is the holonomy of a complete structure on S_i with each boundary curve (as written, $C_1 \ldots C_i$, C_{i+1} and $(C_1 \cdots C_{i+1})^{-1}$) bounding S_i on its right. Each decomposition curve is of the form $(C_1 C_2 \ldots C_j)$ for some j, and appears in two relators, once as itself and once as an inverse. Hence in the corresponding fundamental domains, the curve corresponding to $(C_1 \ldots C_l)$ bounds one fundamental domain on its right, and the other on its left. So there is no folding; orientations will work out correctly, provided we can piece the fundamental domains together. Note that although our fundamental domains may be degenerate along some edges, corresponding to parabolic boundary components, the decomposition curves are all hyperbolic, and so these edges are *not* degenerate.

Note also that fundamental domains might not piece together particularly nicely: for instance, in an octagonal fundamental domain for a pair of pants S_i, one of the boundary curves corresponds not to one but to *two* sides of the octagon. So we will instead piece together *developing maps.*

Consider our first piece, say S_1, with $q = q_1$, and a preferred lift $\tilde{q} = \tilde{q}_1$ in the universal cover $\tilde{S}_1$. Using our construction we obtain an octagonal fundamental domain, with basepoint $\bar{q} = \bar{q}_1 \in \mathbb{H}^2$ and partial lift of $\mathcal{T}$. Hence we have a geometric structure on a surface $W = S_1$ with the following properties:

(i) We have a developing map $\mathcal{D}_W : \tilde{W} \longrightarrow \mathbb{H}^2$ with image the convex core of

$\rho(\pi_1(W))$, giving a hyperbolic structure on W with totally geodesic or cusped boundary components and holonomy ρ_W.

(ii) Suppose C_j is a boundary curve of W which intersects $\mathcal{T}$, i.e. C_j is a decomposition curve. We consider $C_j \in \pi_1(W, q)$ by taking a path from q to the decomposition curve along $\mathcal{T}$; then traversing C_j; and then back to q along $\mathcal{T}$. Using $\mathcal{T}$ in this way we obtain a canonical boundary edge $\tilde{C}_j$ of the universal cover $\tilde{W}$, where $\tilde{C}_j \cong \mathbb{R}$ covers $C_j \cong S^1$. The developing image of $\tilde{C}_j$ is precisely $\text{Axis}\, c_j$ (where $c_j = \rho(C_j)$), which is a boundary edge of the convex core.

Now suppose inductively we have a subsurface W consisting of the union of several adjacent S_i, including S_1. We clearly have an inclusion $\iota_W : \pi_1(W, p) \hookrightarrow \pi_1(S, p)$, and hence a representation $\rho_W = \rho \circ \iota_W$. Suppose the above two conditions are satisfied for W; we will show how to adjoin a new pair of pants S_k adjacent to W and obtain a geometric structure on $W' = S_k \cup W$ with the same properties.

Assuming that we have chosen a lift $\tilde{q} \in \tilde{\mathcal{T}}$ of $q \in \mathcal{T}$ in $\tilde{S}$, there are canonical inclusions $\tilde{S}_k \hookrightarrow \tilde{W}' \hookrightarrow \tilde{S}$ and $\tilde{W} \hookrightarrow \tilde{W}' \hookrightarrow \tilde{S}$. Now $S_k \cap W$ is a single decomposition curve, say C_k. The curve C_k has preferred lifts into $\tilde{S}_k$ and $\tilde{W}$, using $\tilde{\mathcal{T}}$; it is a boundary edge in both. But when we consider the inclusions into $\tilde{W}'$, we see that these two preferred lifts of C_k agree, since there is only one lift of C_k in $\tilde{W}'$ intersecting $\tilde{\mathcal{T}}$. So within $\tilde{W}'$, the universal covers $\tilde{S}_k$ and $\tilde{W}$ intersect precisely along the preferred lift $\tilde{C}_k$. We also obtain a preferred lift $\tilde{q}_k$ of q_k.

The representation ρ_k is the holonomy of a complete hyperbolic structure on S_k. Recall from the discussion of section 8.2.1 that $\rho \circ \zeta_k$ is the representation ρ with basepoint considered to be q_k, with lift at $\tilde{q}_k$. And ρ_k is defined as the composition $(\rho \circ \zeta_k) \circ \iota_k$. So ρ_k is the holonomy of a complete hyperbolic structure on S_k, with basepoint q_k (considered in $S_k \hookrightarrow W'$) lifting to $\tilde{q}_k \in \tilde{S}_k \hookrightarrow \tilde{W}'$. Write $C_k \in \pi_1(S, q)$ and $\hat{C}_k \in \pi_1(S_k, q_k)$ for the elements of the respective fundamental groups corresponding to this boundary element, so $C_k = \zeta_k \circ \iota_k(\hat{C}_k)$. We see that the developing image under $\mathcal{D}_k$ of the common boundary $\tilde{C}_k$ is precisely $\text{Axis}\, \rho_k(\hat{C}_k) = \text{Axis}\, \rho(C_k) = \text{Axis}\, c_k$.

We will adjust $\mathcal{D}_k$ carefully so that $\mathcal{D}_k$ agrees with $\mathcal{D}_W$ along $\tilde{S}_k \cap \tilde{W} = \tilde{C}_k$. In $\mathbb{H}^2$ the developing maps $\mathcal{D}_W : \tilde{W} \to \mathbb{H}^2$ and $\mathcal{D}_k : \tilde{S}_k \to \mathbb{H}^2$ are homeomorphisms and have images which are convex sets intersecting along $\text{Axis}\, c_k$. Both $\mathcal{D}_k$ and $\mathcal{D}_W$ map $\tilde{C}_k$ to $\text{Axis}\, c_k$, and their holonomies agree: $\rho_k(\hat{C}_k) = \rho(C_k) = \rho_W(C_k)$. By our previous comments regarding orientation, one convex set lies on each side of $\text{Axis}\, c_k$. Now there is a diffeomorphism ϕ of $\tilde{C}_k \cong \mathbb{R}$ so that $\mathcal{D}_k|_{\tilde{C}_k} \circ \phi = \mathcal{D}_W|_{\tilde{C}_k}$. This ϕ is

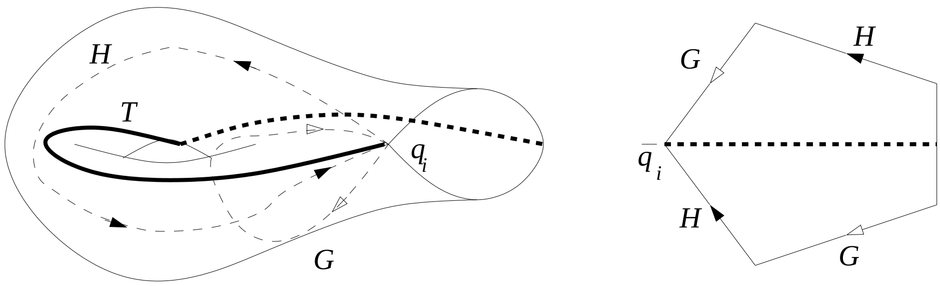

Figure 8.5: Choice of $\mathcal{T}$ on each punctured torus, and developing image.

periodic with period equal to the translation distance of c_k, hence descends to a diffeomorphism of the boundary curve $C_k \subset S_k$, and then extends to a diffeomorphism of S_k. Thus ϕ extends to a diffeomorphism of $\tilde{S}_k$ so that $\mathcal{D}_k \circ \phi$ is a developing map for a complete hyperbolic structure on S_k with all the usual properties, and which agrees with $\mathcal{D}_W$ along the common boundary $\tilde{C}_k$.

Combining the two developing maps $\mathcal{D}_k, \mathcal{D}_W$ gives a partial developing map of W', which extends equivariantly to a complete developing map $\mathcal{D}_{W'}$ for W'. The image is the convex core of $\rho(\pi_1(W'))$, and we obtain a hyperbolic structure on W' with totally geodesic boundary. Let the holonomy be ρ'; we will show that $\rho' = \rho$. We know that ρ and ρ' agree on $\pi_1(W)$. But recall that $\rho_k = \rho \circ \zeta_k \circ \iota_k$, and ρ_k is the holonomy of $S_k \hookrightarrow W'$ considered with basepoint q_k lifting to $\tilde{q}_k$. Hence the holonomy ρ' of a curve in $S_k \hookrightarrow W'$ considered with basepoint p is given by ρ. Thus $\rho = \rho'$ on $\pi_1(W)$ and on $\pi_1(S_k)$, hence on $\pi_1(W')$.

Now we have verified all the required properties on W'. So continuing in this manner, we obtain a hyperbolic structure on the entire surface S, where S has genus 0.

Now consider the general case. We can consider our surface S as an $(n+g)$-holed sphere with g punctured tori attached. So supposing we have a surface W satisfying the two properties above, we only need to show that we can attach a punctured torus S_k, obtaining a new surface W' with the same two properties.

The same argument applies, but we will be careful with our choice of $\mathcal{T}$. For a given basis G_k, H_k of $\pi_1(S_k, q_k)$, choose the map $\mathcal{T}$ of the dual tree to run between G_i, H_i as shown in figure 8.5. We do this so that in a developing map, $\mathcal{T}$ remains inside the fundamental domain, running direct from basepoint to boundary.

We immediately obtain a hyperbolic structure on S_k with holonomy ρ_k. We let $C_k = S_k \cap W$. Again C_k has a preferred lift $\tilde{C}_k$ in $\tilde{W}'$ which is the intersection of $\tilde{S}_k$ and $\tilde{W}$. The curve $C_k \in \pi_1(S_k, q_k)$ is defined by travelling along $\mathcal{T}$ from q_k to the boundary, then traversing the boundary in the direction of $[G_k, H_k]$, then travelling back to q_k along $\mathcal{T}$. In terms of our basis then, $C_k = G_k^{-1} H_k^{-1} G_k H_k = [G_k^{-1}, H_k^{-1}]$.

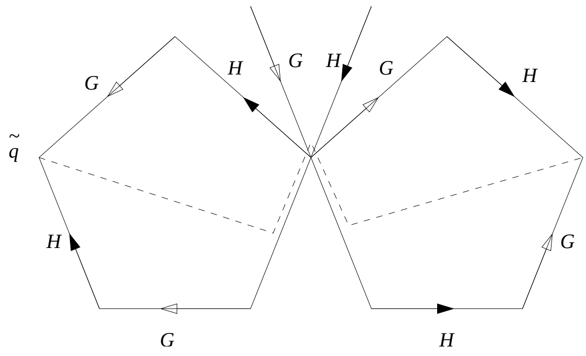

Figure 8.6: As an element of $\pi_1(S_k, q_k)$, $C_k = [G_k^{-1}, H_k^{-1}]$.

See figure 8.6 for an illustration of C_k, in the universal cover. By our construction $\mathcal{D}_k$ maps $\tilde{C}_k$ to Axis$[g_k^{-1}, h_k^{-1}]$, so the above two conditions are satisfied.

As before, we adjust $\mathcal{D}_k$ so that $\mathcal{D}_k, \mathcal{D}_W$ agree along $\tilde{C}_k$. Since $\mathcal{E}(\rho_k)[S_k] = 1$ and the boundary curve has hyperbolic holonomy, by proposition 4.3.4 the representation ρ_k falls into the case of section 7.2, and by proposition 7.2.1 the boundary C_k bounds S_k on its left. We found above that C_k bounds W on its right. So the developing maps combine smoothly without folding to give a partial developing map of W', which we extend to a complete developing map for W'. As before we see that the holonomy of this hyperbolic structure is ρ.

Repeating this process we obtain a geometric structure on S with holonomy ρ. For boundary components with hyperbolic components we clearly obtain geodesic boundary. A boundary component with parabolic holonomy must lie in one of the pants of our decomposition, and according to the construction of section 8.2.2 becomes a cusp. This concludes the proof of Goldman's theorem.

8.3 Constructions for the genus 2 surface

Throughout this section, let S be a closed genus 2 surface. We prove theorem B. Again assume that $\mathcal{E}(\rho)[S] = 1$; if $\mathcal{E}(\rho)[S] = -1$ then the same arguments apply with opposite orientation. We suppose that there is a separating curve C on S such that $\rho(C)$ is not hyperbolic.

8.3.1 Splitting into tori

Consider the separating curve C on S, and take a basepoint q on C. So C splits S into two punctured tori S_0, S_1. A dual tree to the splitting is just an edge with a vertex at either end. We take basepoints $q_0 = q_1 = q$ for S_0, S_1, S respectively. On each punctured torus, let $G_i, H_i \in \pi_1(S_i, p_i)$ be a pair of basis curves, so that $[G_0, H_0]$

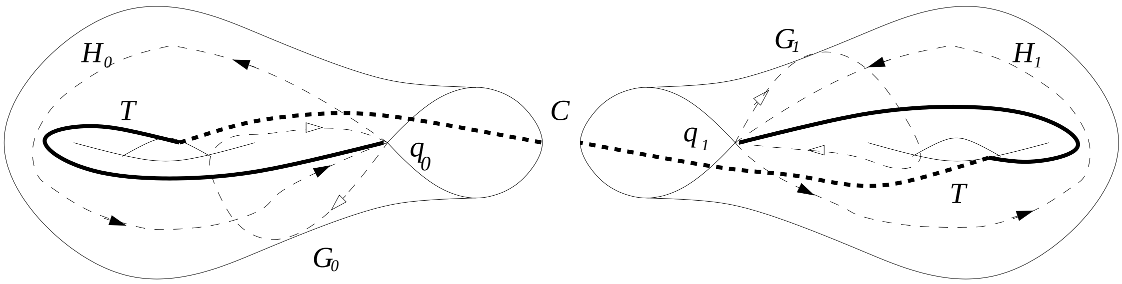

Figure 8.7: Decomposition of S and dual tree $\mathcal{T}$.

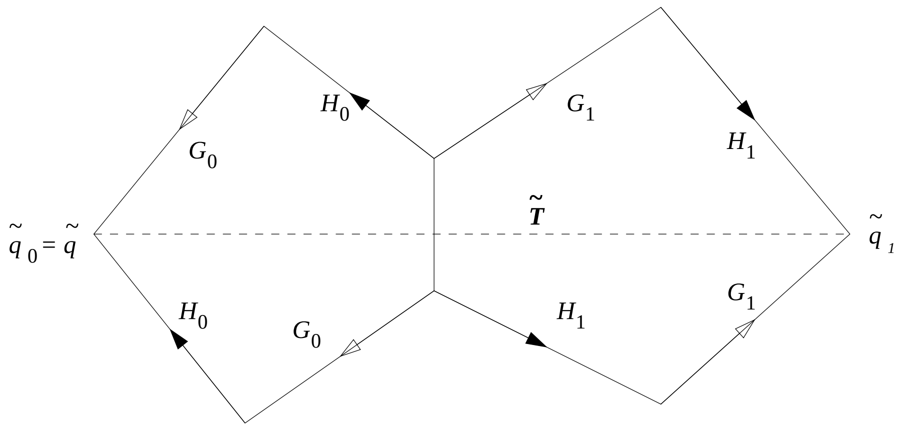

Figure 8.8: Preferred lifts of basepoints and $\mathcal{T}$ in the universal cover $\tilde{S}$.

and $[G_1, H_1]$ are homotopic to C, but traversed in opposite directions. Choose the map $\mathcal{T}$ of the dual tree to run between G_i, H_i as in figure 8.5, as in the previous section. From this we obtain preferred lifts $\tilde{q} = \tilde{q}_0$ and $\tilde{q}_1 \in \tilde{S}$, homomorphisms

$$\iota_0 : \pi_1(S_0, q_0) \hookrightarrow \pi_1(S, q_0) = \pi_1(S, q), \qquad \iota_1 : \pi_1(S_1, q_1) \hookrightarrow \pi_1(S, q_1) = \pi_1(S, q),$$
$$\zeta_0 : \pi_1(S, q_0) \xrightarrow{\text{Id}} \pi_1(S, q), \qquad \zeta_1 : \pi_1(S, q_1) \xrightarrow{\cong} \pi_1(S, q),$$

and representations $\rho_0 = \rho \circ \zeta_0 \circ \iota_0$, $\rho_1 = \rho \circ \zeta_1 \circ \iota_1$. Note by our choice of $\mathcal{T}$, we have $\tilde{q}_0 \neq \tilde{q}_1$, even though $q_0 = q_1$. See figures 8.7 and 8.8.

Let G_0, H_0 be a basis for $\pi_1(S_0, q_0)$ and G_1, H_1 a basis for $\pi_1(S_1, q_1)$. We may certainly choose these bases so that $\pi_1(S, q)$ has the standard presentation

$$\pi_1(S, q) = \langle G_0, H_0, G_1, H_1 \mid [G_0, H_0]\, [G_1, H_1] = 1 \rangle.$$

Let $L \in \pi_1(S, q)$ denote the loop traced out by $\mathcal{T}$ from q_0 to q_1, so that ζ_1 is conjugation by L. We see $L = G_0^{-1} H_0^{-1} G_1 H_1 = H_0^{-1} G_0^{-1} H_1 G_1$ (see figure 8.8). Writing $g_i = \rho(G_i)$, $h_i = \rho(H_i)$, we have

$$\rho_0(G_0) = g_0, \quad \rho_0(H_0) = h_0, \quad \rho_1(G_1) = \rho(L) g_1 \rho(L^{-1}), \quad \rho_1(H_1) = \rho(L) h_1 \rho(L^{-1}).$$

Since $\mathcal{E}(\rho)[S] = 1$ we have by proposition 4.3.1 $[g_0, h_0][g_1, h_1] = \mathbf{z}$. Note that

$$[g_i, h_i], \quad [g_i^{-1}, h_i^{-1}], \quad [\rho_i(G_i)^{-1}, \rho_i(H_i)^{-1}]$$

are all conjugates in the holonomy group, hence lie in the same region of $\widetilde{PSL_2\mathbb{R}}$. In fact, choosing arbitrary lifts $\tilde{g}_0, \tilde{h}_0, \tilde{g}_1, \tilde{h}_1 \in \widetilde{PSL_2\mathbb{R}}$ and $\widetilde{\rho(L)} = \tilde{g}_0^{-1}\tilde{h}_0^{-1}\tilde{g}_1\tilde{h}_1$, recalling lemma 3.3.1, we obtain:

$$\begin{aligned}
\rho_0([G_0^{-1}, H_0^{-1}])\rho_1([G_1^{-1}, H_1^{-1}]) &= [\tilde{g}_0^{-1}, \tilde{h}_0^{-1}]\widetilde{\rho(L)}[\tilde{g}_1^{-1}, \tilde{h}_1^{-1}]\widetilde{\rho(L)}^{-1} \\
&= [\tilde{g}_0^{-1}, \tilde{h}_0^{-1}]\tilde{g}_0^{-1}\tilde{h}_0^{-1}\tilde{g}_1\tilde{h}_1[\tilde{g}_1^{-1}, \tilde{h}_1^{-1}]\tilde{h}_1^{-1}\tilde{g}_1^{-1}\tilde{h}_0\tilde{g}_0 \\
&= \tilde{g}_0^{-1}\tilde{h}_0^{-1}[\tilde{g}_0, \tilde{h}_0][\tilde{g}_1, \tilde{h}_1]\tilde{h}_0\tilde{g}_0 \\
&= \tilde{g}_0^{-1}\tilde{h}_0^{-1}\mathbf{z}\tilde{h}_0\tilde{g}_0 = \mathbf{z}.
\end{aligned}$$

As $[G_i, H_i]$ is homotopic to C, traversed in some direction, we have each $[g_i, h_i]$ is not hyperbolic. From this and proposition 3.7.2 we have

$$[g_0, h_0], [g_1, h_1] \in \{1\} \cup \text{Ell}_{-1} \cup \text{Ell}_1 \cup \text{Par}_{-1}^+ \cup \text{Par}_0 \cup \text{Par}_1^- .$$

As $[g_0, h_0], [g_1, h_1]$ are inverses in $PSL_2\mathbb{R}$, they are both elliptic, both parabolic, or both the identity. Applying properties of Θ discussed in section 3.5, we have $\Theta([g_0, h_0]) + \Theta([g_1, h_1]) = \pi$. From the bounds in corollary 3.6.4, and assuming without loss of generality that $\Theta([g_0, h_0]) \leq \Theta([g_1, h_1])$, we see there are only the following two possibilities.

(i) **Elliptic case.** $[g_0, h_0] \in \text{Ell}_1$ with $\Theta([g_0, h_0]) \in (0, \pi/2]$, and $[g_1, h_1] \in \text{Ell}_1$ with $\Theta([g_1, h_1]) \in [\pi/2, \pi)$.

(ii) **Parabolic case.** $[g_0, h_0] \in \text{Par}_0^+$ and $[g_1, h_1] \in \text{Par}_1^-$.

Note in particular that $[g_0, h_0], [g_1, h_1]$, considered as elements of $PSL_2\mathbb{R}$, cannot be the identity. We will consider these two cases separately in the next two sections.

8.3.2 Piecing together along an elliptic

We have $[g_0, h_0], [g_1, h_1] \in \text{Ell}_1$ with $\Theta([g_0, h_0]) \in (0, \pi/2]$ and $\Theta([g_1, h_1]) \in [\pi/2, \pi)$. We will rely heavily on proposition 7.4.1. We see that ρ_0 is the holonomy of a hyperbolic cone-manifold structure on S_0 with no interior cone points and at most one corner point, with corner angle in $[2\pi, 3\pi)$: this is the "large angle" elliptic case. Similarly, ρ_1 is a "small angle" elliptic case and is the holonomy of a cone-manifold structure with corner angle in $(\pi, 2\pi]$.

Recall how these constructions work, from section 7.3. We construct a pentagonal fundamental domain by choosing a point p close to the fixed point of $[g^{-1}, h^{-1}]$, and take $\mathcal{P}(g, h, p)$. By choosing the point $p \in \mathbb{H}^2$ judiciously, the pentagon $\mathcal{P}(g, h, p)$ is non-degenerate and bounds an embedded disc, giving rise to the desired cone-manifold structure. Since $[g_0^{-1}, h_0^{-1}]$ and $[g_1^{-1}, h_1^{-1}]$ are both in Ell_1, by proposition 7.4.1, $[G_0, H_0]$ will bound S_0 on its left and $[G_1, H_1]$ will bound S_1 on its left. Since $[G_0, H_0]$ and $[G_1, H_1]$ are homotopic to C traversed in opposite directions, if we can piece the fundamental domains together along a common boundary, there will be no folding.

The representation $\rho_0 : \pi_1(S_0, q_0) \to PSL_2\mathbb{R}$ maps $(G_0, H_0) \mapsto (g_0, h_0)$. We take p_0 close to the fixed point of $\rho_0([G_0^{-1}, H_0^{-1}]) = [g_0^{-1}, h_0^{-1}]$ and consider $\mathcal{P}(g_0, h_0, p_0)$. This is a fundamental domain for a developing map for a hyperbolic cone-manifold structure on S_0. We choose a preferred lift $\tilde{q}_0 \in \tilde{S}_0 \hookrightarrow \tilde{S}$ of q_0. Then we have preferred lifts $\tilde{\mathcal{T}}$ and $\tilde{q}_1$. By our choice of $\mathcal{T}$, the points $\tilde{q}_0$ and $\tilde{q}_1$ lie on adjacent pentagonal fundamental domains as in figure 8.8.

We start with this fundamental domain, as in the proof of Goldman's theorem, and add on the fundamental domain for S_1. The representation $\rho_1 : \pi_1(S_1, q_1) \to PSL_2\mathbb{R}$ is given by $\rho \circ \zeta_1 \circ \iota_1$ and will correspond to the holonomy of a developing map where q_1 lifts to $\tilde{q}_1 \in \tilde{S}_1 \hookrightarrow \tilde{S}$. We construct $\mathcal{P}(\rho_1(G_1), \rho_1(H_1), p_1)$ where p_1 is close to the fixed point of $[\rho_1(G_1)^{-1}, \rho_1(H_1)^{-1}]$, which is the same as the fixed point of its inverse $[\rho_0(G_0)^{-1}, \rho_0(H_0)^{-1}] = [g_0^{-1}, h_0^{-1}]$.

If we can choose basepoints p_0, p_1 such that the pentagons $\mathcal{P}(\rho_0(G_0), \rho_0(H_0), p_0)$ and $\mathcal{P}(\rho_1(G_1), \rho_1(H_1), p_1)$ join precisely along the edges representing their boundary, without folding, then we will have an immersed non-degenerate geodesic octagon bounding an immersed disc in $\mathbb{H}^2$. The pentagonal fundamental domains piece together in $\mathbb{H}^2$ just as they do in $\tilde{S}$, so we obtain a fundamental domain of a developing map for a cone-manifold structure on S, and we may then extend equivariantly to a complete developing map, obtaining a cone-manifold structure on S. We must have $p_1 = [g_0^{-1}, h_0^{-1}]p_0$, as shown in figure 8.9. There will be at most one cone point, given by the vertices (all of which are identified) in the fundamental domain. The angle at the cone point will be the sum of the interior angles of the octagon, which is equal to the sum of the two corner angles in the two punctured tori.

Recall now proposition 7.4.1. Let r denote the fixed point of $[g_0^{-1}, h_0^{-1}]$. There exists a closed semicircular disc $C_{\epsilon_0}(r)$ centred at r such that if p_0 is chosen anywhere in this disc (except r), then $\mathcal{P}(\rho_0(G_0), \rho_0(H_0), p_0)$ is non-degenerate and bounds an embedded disc. Similarly there exists a semicircular disc $C_{\epsilon_1}(r)$ for which $\mathcal{P}(\rho_1(G_1), \rho_1(H_1), p_1)$ is non-degenerate and bounds an embedded disc. Take $\epsilon = \min(\epsilon_1, \epsilon_2)$. Note that on the circle of radius ϵ about q, there is a closed arc of angle π on which p_0 can be validly chosen. Hence there is a closed arc of angle π

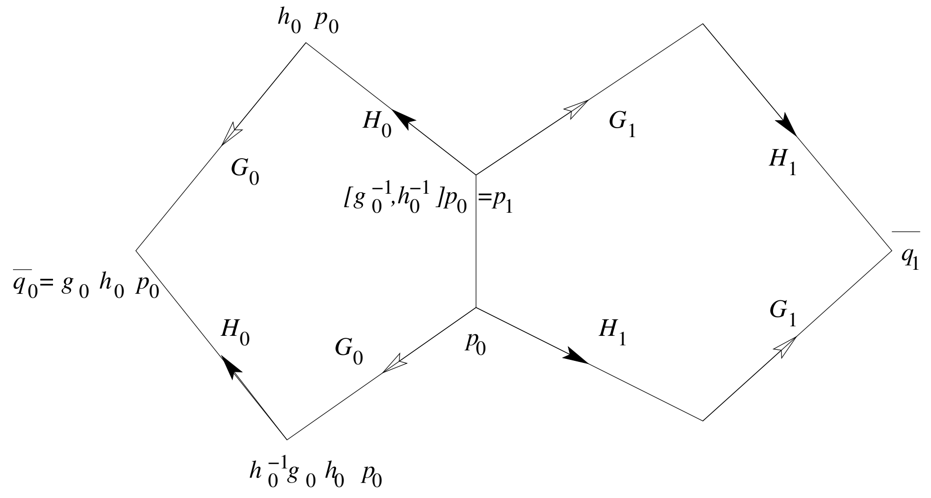

Figure 8.9: Putting the pieces together.

on which $[g_0^{-1}, h_0^{-1}]p_0$ can validly lie. Similarly there is a closed arc of angle π on which p_1 can be validly chosen. Since there is only 2π worth of angle in a circle, these arcs must overlap, and hence we may take p_0 and $p_1 = [g_0^{-1}, h_0^{-1}]p_0$ so that both $\mathcal{P}(\rho_0(G_0), \rho_0(H_0), p_0)$ and $\mathcal{P}(\rho_1(G_1), \rho_1(H_1), p_1)$ are non-degenerate and bound embedded discs. Since $p_1 = [g_0^{-1}, h_0^{-1}]p_0$, the boundary edges of the pentagons are both geodesic segments $p_0 \to p_1$, and since the pentagons have the same orientation, there is no folding. So we obtain a hyperbolic cone-manifold structure on S. As in the proof of Goldman's theorem, the holonomy is ρ.

The cone angle is easy to compute, since again by proposition 7.4.1, $\theta = 3\pi - \text{Tw}([g^{-1}, h^{-1}], p_i)$. Since $[g_0, h_0][g_1, h_1] = \mathbf{z}$, the conjugates $[g_0^{-1}, h_0^{-1}]$, $\rho(L)[g_1^{-1}, h_1^{-1}]\rho(L)^{-1}$ (which are also inverses) multiply to $\mathbf{z}$ also. So we have $\text{Tw}([\rho_0(G_0), \rho_0(H_0)], p_0) + \text{Tw}([\rho_1(G_1), \rho_1(H_1)], p_1) = 2\pi$. Hence the cone angle is $6\pi - 2\pi = 4\pi$, as desired.

8.3.3 Piecing together along a parabolic

We have $[g_0, h_0] \in \text{Par}_0^+$ and $[g_1, h_1] \in \text{Par}_1^-$, so by proposition 3.7.1 we have $\text{Tr}[g_0, h_0] = 2$, $\text{Tr}[g_1, h_1] = -2$. Hence we may apply the results of sections 8.5 and 8.3 respectively. The strategy is the similar to the previous section. Let $r = \text{Fix}[\rho_0(G_0)^{-1}, \rho_0(H_0)^{-1}] = \text{Fix}[\rho_1(G_1)^{-1}, \rho_1(H_1)^{-1}]$.

First consider S_0. From the discussion in section 8.5, we may take a basis G_0, H_0 of $\pi_1(T_1)$ and a point p_0 arbitrarily close to r such that $\mathcal{P}(\rho_0(G_0), \rho_0(H_0), p_0)$ is non-degenerate and bounds an embedded disc. Since $[\rho_0(G_0), \rho_0(H_0)] = [g_0, h_0] \in \text{Par}_0^+$, so ∂S_0 traversed in the direction of $[G_0, H_0]$ bounds S_0 on its left. This gives a hyperbolic cone-manifold structure on S_0 corresponding to a preferred lift $\tilde{q}_0$ of q_0 with holonomy ρ_0. The corner angle is $3\pi - \text{Tw}([\rho_0(G_0)^{-1}, \rho_0(H_0)^{-1}]$.

Now consider S_1. As above, we take a basis G_1, H_1 of $\pi_1(S_1, q_1) \overset{\iota_1}{\hookrightarrow} \pi_1(S, q)$ such that $[G_0, H_0][G_1, H_1] = 1$. By the discussion in section 8.3, the representation ρ_1 is discrete, and the quotient of $\mathbb{H}^2$ by the image of ρ_1 is a cusped torus. We may take p_1 anywhere sufficiently close to r, and obtain $\mathcal{P}(\rho_1(G_1), \rho_1(H_1), p_1)$ non-degenerate bounding an embedded disc. This gives a hyperbolic cone-manifold structure on S_1 corresponding to the preferred lift $\tilde{q}_1$ of q_1 with holonomy ρ_1. The corner angle $3\pi - \text{Tw}([\rho_1(G_1)^{-1}, \rho_1(H_1)^{-1}], p)$, and boundary ∂S_1 traversed in the direction of $[G_1, H_1]$ bounding S_1 on its left.

Hence we may take p_0, p_1 such that $p_1 = [g_0^{-1}, h_0^{-1}]p_0$ and both $\mathcal{P}(\rho_0(G_0), \rho_0(H_0), p_0)$, $\mathcal{P}(\rho_1(G_1), \rho_1(H_1), p_1)$ are non-degenerate pentagons bounding immersed discs. Since they both have the same orientation, they fit together without folding along their boundary edges to give a non-degenerate octagon in $\mathbb{H}^2$ bounding an immersed disc, and hence a cone-manifold structure on S_2. And again the holonomy is ρ. Since $[\rho_0(G_0)^{-1}, \rho_0(H_0)^{-1}][\rho_1(G_1)^{-1}, \rho_1(H_1)^{-1}] = \mathbf{z}$ we have $\text{Tw}([\rho_0(G_0)^{-1}, \rho_0(H_0)^{-1}], p_0) + \text{Tw}([\rho_1(G_1)^{-1}, \rho_1(H_1)^{-1}], p_1) = 2\pi$. Hence the cone angle is $6\pi - 2\pi = 4\pi$.

Geometrically, one half of S has the nice structure of a truncated cusped torus, and the other half is a rather uglier handle tacked on to the truncated cusp. This concludes the proof of theorem B.

8.4 Representations with $\mathcal{E}(\rho)[S] = \pm(\chi(S) + 1)$

8.4.1 Simple cases

We turn now to the situation where S is a closed surface of genus $g \geq 2$ and ρ is a representation with $\mathcal{E}(\rho)[S] = \pm(\chi(S) + 1)$. Recall the statement of theorem C assumes that there exists a non-separating simple closed curve C with $\rho(C)$ elliptic.

Note that such representations exist: theorem C is not vacuous. We know that there exist representations of the fundamental group of a punctured torus S_1 which take a non-separating simple closed curve to an elliptic and have $\text{Tr}[g, h] > 2$: there are points $(x, y, z) \in X(S_1)$ with $|x| < 2$ and $\kappa(x, y, z) > 2$. See theorem 6.1.2 and figure 6.1. Consider S as constructed by pasting together S_1 and a surface W of genus $g - 1$ with one boundary component. We may take ρ as above on S_1; which is the holonomy of a complete hyperbolic structure on W; and which agrees along the common boundary curve. Such a representation has the desired properties. By proposition 4.2.2, proposition 4.3.4 and lemma 4.2.1 we have $\mathcal{E}(\rho)[S] = 0 + (\pm\chi(W)) = \pm(\chi(S) + 1)$.

Given the non-separating simple closed curve C, we can find a separating curve D, disjoint from C, cutting S into two pieces, and so that the side containing C

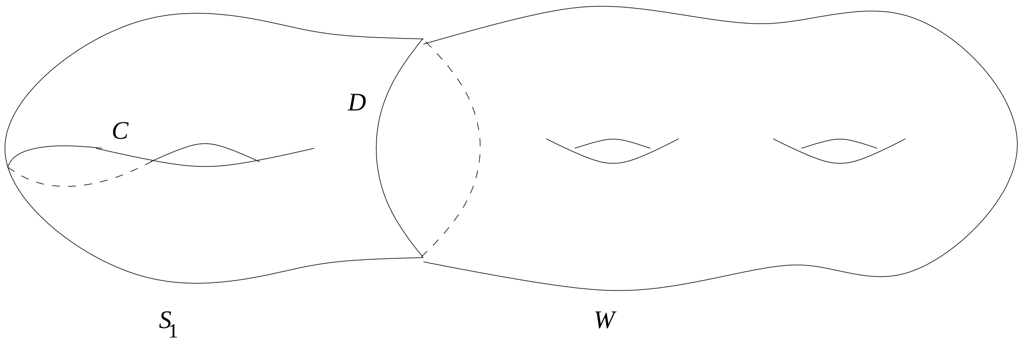

Figure 8.10: Decomposition of S with an elliptic on one side.

is a punctured torus S_1. Let the surface obtained on the other side be W, so W has genus $g-1$ and 1 boundary component. Choosing basepoints $q = q_W$ for S, W and q_1 for S_1, and a dual tree $\mathcal{T}$ as described in sections 8.2 and 8.3, we obtain representations ρ_W, ρ_1 on W and S_1.

We may now take a basis G, H for $\pi_1(S_1, q_1)$ with G homotopic to C, and with $[G, H]$ homotopic to D. So $\rho_1(G) = g$ is elliptic. Recall from lemma 3.2.2 that if $\text{Tr}[g, h] < 2$ then g, h are hyperbolic; so $\text{Tr}[g, h] \geq 2$. Hence $[g, h]$ is not elliptic, and relative Euler classes are well-defined. By proposition 4.3.4, $\mathcal{E}(\rho_1)[S_1] = 0$, so by additivity of the relative Euler class 4.2.1, $\mathcal{E}(\rho_W)[W] = -\chi(W)$. Hence by Goldman's theorem ρ_W is the holonomy of a complete hyperbolic structure on W with totally geodesic or cusped boundary. We will deal with the three cases in turn: $[g, h]$ is the identity, parabolic or hyperbolic. We first rule out the identity and parabolic cases, before focusing on the hyperbolic case.

First suppose $[g, h] = 1$. Then ρ_W takes the boundary component of W to the identity. But as ρ_W is the holonomy of a complete hyperbolic structure on W with totally geodesic or cusped boundary, the holonomy of each boundary curve must be hyperbolic or parabolic. This is a contradiction.

Lemma 8.4.1 *If $\rho(D)$ is parabolic then ρ is the holonomy of a hyperbolic cone-manifold structure on S with one cone point of angle 4π.*

Proof We use a similar method to section 8.3.3. We already have a complete hyperbolic structure on W, with the boundary curve D corresponding to a cusp. The representation ρ_1 has $\text{Tr}[g, h] = 2$, so ρ_1 is reducible, and non-abelian, hence by proposition 7.5.2 is the holonomy of a hyperbolic cone-manifold structure on S_1 with no cone points and one corner point with corner angle $> 2\pi$. The pentagonal fundamental domain $\mathcal{P}(g, h, p)$ can be chosen with p arbitrarily close to the fixed point r at infinity of $[g^{-1}, h^{-1}]$.

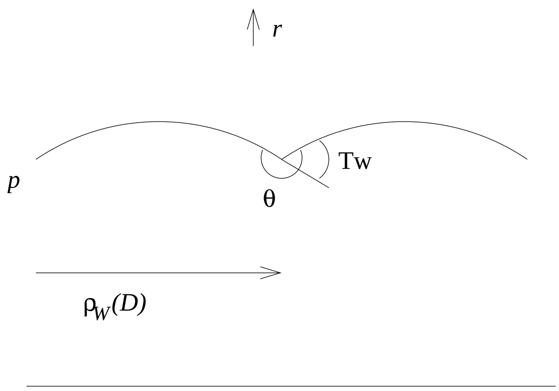

Figure 8.11: Developing image of $\tilde{D}$ under $\mathcal{D}_W$.

Take a developing map $\mathcal{D}_W : \tilde{W} \longrightarrow \mathbb{H}^2$ for W. The boundary curve D has a preferred lift $\tilde{D}$ arising from a choice of basepoint and lift of the tree $\mathcal{T}$ in $\tilde{W}$. The universal covers $\tilde{S}_1$ and $\tilde{W}$ include into $\tilde{S}$ accordingly and $\tilde{S}_1 \cap \tilde{W} = \tilde{D}$. Under $\mathcal{D}_W$, $\tilde{D}$ maps to the fixed point r of $\rho_W(D)$. We can truncate this geometric structure along some curve homotopic to D. This gives a hyperbolic cone-manifold structure on W with no cone points and one corner point. With this truncated developing map, $\tilde{D}$ maps to a piecewise geodesic curve close to r. By choosing the point p described above sufficiently close to r, we may truncate W so that in the developing map, p corresponds to one of the developing images of the corner point. Taking a simplest lift of $\rho_W(D)$ in $\widetilde{PSL_2}\mathbb{R}$, the corner angle in W is then $\pi + \text{Tw}(\rho_W(D), p)$: see figure 8.11. By adjusting W through a diffeomorphism supported near D, we may assume that the geodesic segment correpsonding to the boundary edge of $\mathcal{P}(g, h, p)$, and the developing image of $\tilde{D}$ under $\mathcal{D}_W$, agree. Thus we have partial developing maps for W and S_1 which piece together as in the proof of Goldman's theorem. These extend equivariantly to give us a hyperbolic cone-manifold structure on S. There is no folding: as in the proof of Goldman's theorem, if we orient curves appropriately then as $\mathcal{E}(\rho_W)[W] = -\chi(T)$, under $\mathcal{D}_W$ we see $\tilde{D}$ bounds $\tilde{W}$ on its right, while in the fundamental domain for S_1 since $\mathcal{E}(\rho_1)[S_1] = 1$, we see $\tilde{D}$ bounds the pentagon on its left. By proposition 7.5.2 the corner angle in S_1 is given by $3\pi - \text{Tw}(\rho_1(D), p)$. Thus the two corner angles piece together to give a single cone point with angle 4π. And by a similar argument to Goldman's theorem, the holonomy is given by ρ. ∎

We are now left only with the case where $\rho(D)$ is hyperbolic, i.e. $\text{Tr}[g, h] > 2$.

8.4.2 Piecing together along a hyperbolic

We have a closed surface S of genus ≥ 2. We have cut S into two pieces along a curve D, obtaining a punctured torus S_1 and another piece W, with representations

ρ_1, ρ_W. We have a basis G, H for $\pi_1(S)$ and $g = \rho_1(G)$ elliptic. We may now assume that $\text{Tr}[g,h] > 2$. The relative Euler classes are given by $\mathcal{E}(\rho_1)[S_1] = 0$ and $\mathcal{E}(\rho_T)[W] = -\chi(W)$. The representation ρ_W is a holonomy representation for a complete hyperbolic structure on W, and is therefore discrete.

As ρ_W is discrete, the quotient of $\mathbb{H}^2$ by the image of ρ_W is a flared surface. For a surface with totally geodesic boundary, we truncate the "flares" along geodesics in the homotopy classes of the boundary curves. But, as described in chapter 5, we may also truncate a "flare" away from the geodesic, and obtain a piecewise geodesic boundary, with a single corner point. We may truncate arbitrarily anywhere at all "outside" the geodesic, obtaining corner angles in $(0, \pi)$. Further, as described in chapter 5, we may truncate "inside" the geodesic, producing a corner angle in $(\pi, 2\pi)$. We cannot truncate too far inside the surface, but if we stay within the collar width of the geodesic then we are guaranteed still to obtain a cone-manifold structure on W. The collar width w is given by

$$\sinh w = \frac{1}{\sinh\left(\frac{d_D}{2}\right)};$$

where d_D is the length of the geodesic corresponding to the boundary curve D. See [10] for details.

We wish to perform such a truncation, to find a hyperbolic cone-manifold structure on W which pieces together with that on S_1 to give a cone-manifold structure on S, in a similar manner to the parabolic case above. If we can find a pentagonal fundamental domain for S_1, which pieces together with the developing map $\mathcal{D}_W$ along the developing image of $\tilde{D}$, then (possibly after adjusting $\mathcal{D}_W$ by precomposing with some homeomorphism of W) we will have a partial developing map for a hyperbolic cone-manifold structure on S with one cone point. As in the parabolic case, there will be no folding. The corner angle on S_1 will be given by $3\pi - \text{Tw}(\widetilde{\rho(D)}, p)$, by section 7.6, since $[g,h] \in \text{Hyp}_0$, taking the simplest lift of $\rho(D)$ in $\widetilde{PSL_2\mathbb{R}}$. The corner angle on W will be given by $\pi + \text{Tw}(\rho(D), p)$. Here we take the simplest lift of $\rho(D)$ into $\widetilde{PSL_2\mathbb{R}}$. So as in the parabolic case the cone angle will automatically be 4π.

These considerations essentially reduce the problem to the following two results, to which the next two sections are dedicated.

Proposition 8.4.2 *Let (G, H) be a basis for $\pi_1(S_1)$. Let $t > 2$, and let $X_t(S_1) = \kappa^{-1}(t) \cap X(S_1)$ denote the relative character variety. Let $\Omega_t \subset X_t(S_1)$ be the set of characters of representations $\rho : \pi_1(S_1) \longrightarrow PSL_2\mathbb{R}$ such that $\rho(C)$ is elliptic for some simple closed curve C on S_1. For almost every ρ with character in Ω_t, the following statement is true: for any $\epsilon > 0$, there exists a basis (G', H') for*

$\pi_1(S_1)$, of the same orientation as (G, H), and a point p at distance less than ϵ from $\mathrm{Axis}[g'^{-1}, h'^{-1}]$*, such that the pentagon $\mathcal{P}(g', h', p)$ is non-degenerate, bounds an embedded disc, and is of a specified orientation.*

Recall the relative character variety $X_t(S_1)$ has a symplectic structure given by the form ω described in 6.3, and hence has a measure μ_t given by integrating ω. The proof of this result will use ergodicity properties of the action on the character variety. The strategy is to show that *some* representation ρ^* (not necessarily our given ρ) of the class considered produces such a pentagon; and to use ergodicity to show that changing basis we can "almost" move anywhere within the character variety, so we can get close to this representation and produce such a pentagon.

Note that the word "almost" cannot be removed from the statement of 8.4.2: virtually abelian representations certainly lie inside the class of representations considered, and the conclusion is certainly false in this case.

Call the set of representations for which the statement is false $B_t \subset X_t(S_1)$, "bad" representations. Proposition 8.4.2 says that $\mu_t(B_t) = 0$.

Let $\mathcal{U}$ denote the set of separating curves D which split S into a punctured torus S_1 and another surface W. For $D \in \mathcal{U}$ and $t > 2$, let $B_{D,t} \subset X(S)$ denote the set of all characters of representations ρ such that

(i) $\mathcal{E}(\rho)[S] = -\chi(S) - 1$;

(ii) ρ takes some simple closed curve on S_1 to an elliptic (hence $\mathcal{E}(\rho_1)[S_1] = 0$ and $\mathcal{E}(\rho_W)[W] = -\chi(W)$);

(iii) with respect to some dual tree $\mathcal{T}$, the induced representation ρ_1 is bad.

(iv) with respect to some basis G, H of $\pi_1(S_1)$, $\mathrm{Tr}[g, h] = t > 2$

(Note there are many possible induced representations ρ_1, since there are many choices of dual trees; but all such induced representations are conjugate, so the above statement makes sense. Clearly ρ_1 is bad iff any conjugate is bad. And $\mathrm{Tr}[g, h]$ does not depend on any choices either.) Let $B_D = \cup_t B_{D,t}$ and $B = \cup_D B_D$.

Proposition 8.4.3 $\mu_S(B) = 0$.

PROOF (OF THEOREM C ASSUMING 8.4.2 AND 8.4.3) (Mostly this is a summary of preceding arguments.)

By proposition 8.4.3, it suffices to show that if ρ is a representation with character in $Y \backslash B$ then ρ is the holonomy of a cone-manifold structure on S with a single cone point with cone angle 4π. Let ρ be such a representation, so $\mathcal{E}(\rho)[S] = \pm(\chi(S) + 1)$;

clearly it is sufficient to prove the result in the case $\mathcal{E}(\rho)[S] = -\chi(S) - 1$, and the other case is identical with reversed orientation.

As ρ has character in Y, there is a non-separating simple close curve C with $\rho(C)$ elliptic. We cut along a separating simple closed curve D to obtain a punctured torus S_1 containing C, and another surface W. From the argument in section 8.4.1, $\rho(D)$ is parabolic or hyperbolic. If $\rho(D)$ is parabolic then by lemma 8.4.1 ρ is a holonomy representation as desired, so assume $\rho(D)$ is hyperbolic. Choose a basepoint $q = q_1$ for both S and S_1 on ∂S_1, and a basis G, H for $\pi_1(S)$, where G is homotopic to C. We thus obtain a representation ρ_1 and write $g = \rho(G)$, $h = \rho(H)$, and g is elliptic. As discussed in section 8.4.1, $\mathrm{Tr}[g,h] = t > 2$ and $\mathcal{E}(\rho_1)[S_1] = 0$.

Now the character of ρ is not in B, hence not in $B_{D,t}$. So ρ_1 has character in Ω_t, but not in $B_t \subset X_t(S_1)$. Hence by proposition 8.4.2, we may take another basis G', H' of $\pi_1(S, q_1)$, of the same orientation, and p within the collar width of $\mathrm{Axis}[g'^{-1}, h'^{-1}]$, such that $\mathcal{P}(g', h', p)$ is non-degenerate, bounds an embedded disc, and the edge $p \to [g'^{-1}, h'^{-1}]p$ bounds the pentagon on its left.

Now $[G'^{-1}, H'^{-1}] = D'$ (as in section 8.2.3 and figure 8.6) is conjugate to $[G^{-1}, H^{-1}] = D$ in $\pi_1(S_1)$, by Nielsen's theorem 6.2.1, as it has the same orientation. Given this basis (G', H'), choose a dual tree $\mathcal{T}$ which on S_1 is as discussed in section 8.2.3 and illustrated in figure 8.5. Choose a preferred lift $\tilde{q} = \tilde{q}_1$ of the basepoint $q = q_1$, and hence preferred lifts $\tilde{\mathcal{T}}, \tilde{q}_W$, $\tilde{D}'$ and $\tilde{W}$. Then $\tilde{D}'$ is the common boundary of the lifts $\tilde{S}_1, \tilde{W} \subset \tilde{S}$. The pentagon $\mathcal{P}(g', h', p)$ is the fundamental domain for a hyperbolic cone-manifold structure on S_1, with boundary edge close to $\mathrm{Axis}(\rho(D')) = \mathrm{Axis}[g'^{-1}, h'^{-1}]$, and holonomy ρ_1.

The induced representation ρ_W on W has $\mathcal{E}(\rho_W)[W] = -\chi(W)$, so is the holonomy of a complete structure on W. The developing map $\mathcal{D}_W$ takes $\tilde{D}'$ to $\mathrm{Axis}(\rho(D'))$. By truncating the surface W slightly, and applying a homeomorphism supported near ∂W, we obtain developing maps which piece together to give a partial developing map for S. There is no folding, as discussed above, and one cone point of angle 4π. As seen in previous arguments, the holonomy is ρ. So we have the desired geometric structure. ∎

8.4.3 Ergodicity

Let $\rho : \pi_1(S_1) \longrightarrow PSL_2\mathbb{R}$ be a representation and let G, H be a basis of $\pi_1(S_1)$. Lifting g, h to $SL_2\mathbb{R}$ arbitrarily and using the notation of chapter 6, let $(\mathrm{Tr}\, g, \mathrm{Tr}\, h, \mathrm{Tr}\, gh) = (x, y, z) \in X(S_1)$. We have already studied the action of $\mathrm{Aut}\, \pi_1(S)$, $\mathrm{Out}\, \pi_1(S_1)$, and

$$\Gamma \cong PGL_2\mathbb{Z} \ltimes \left(\frac{\mathbb{Z}}{2} \oplus \frac{\mathbb{Z}}{2} \right)$$

on triples $(x, y, z) \in X(S_1)$. Recall this action preserves the level sets $X_t(S_1) = \kappa^{-1}(t) \cap X(S_1)$. For the proof of 8.4.2 we are only interested in $t > 2$. Recall that the action of $\operatorname{Out} \pi_1(S_1)$ preserves a symplectic form ω and hence a measure μ_t on each $X_t(S_1)$.

In the case $t > 2$ there are no reducible representations and a character correpsonds uniquely to a conjugacy class of representations. Goldman [30] proved that $X_t(S_1)$ consists of two types of representations:

(i) *Pants representations:* Those $(x, y, z) \in X_t(S_1)$ equivalent to triples (x', y', z') where $x', y', z' \leq -2$. As seen in section 8.2.2, these are discrete representations which can be considered the holonomy of a complete hyperbolic structure on a pair of pants with totally geodesic or cusped boundary. (Note that a given basis will not usually correspond to the boundary components of the pants.) A developing map for this hyperbolic structure tessellates a convex subset of $\mathbb{H}^2$ by non-overlapping fundamental domains, as we saw in section 8.2.2. Thus any element of $\pi_1(S_1)$ has holonomy corresponding to the translation taking one fundamental domain to another. In particular, there are no elliptic elements in the image of ρ; and for any $(x', y', z') \sim (x, y, z)$ we have $|x'|, |y'|, |z'| \geq 2$. We denote the set of such $(x, y, z) \in X_t(S_1)$ by Ψ_t.

(ii) *Representations with elliptics:* Those $(x, y, z) \in X_t(S_1)$ equivalent to (x', y', z') with some coordinate in $(-2, 2)$. That is, there is some simple closed curve on S_1 with elliptic image: we have denote these Ω_t. We have

$$\Omega_t = \Gamma \cdot \left(X_t(S_1) \cap \left[\mathbb{R}^3 \backslash \Big((-\infty, -2] \cup [2, \infty) \Big)^3 \right] \right)$$

Goldman gives an algorithm to change basis and reduce traces until they are small or all negative — a greedy algorithnm which is essentially the opposite of our algorithm from section 7.6.2. Note by definition the action of Γ, or of $\operatorname{Out} \pi_1(S)$, preserves Ω_t and Ψ_t.

Theorem 8.4.4 (Goldman [30]) *For $t > 2$, the action of Γ on Ω_t is ergodic.* ∎

Recall *ergodic* means that the only invariant sets in Ω_t under the action of Γ are null or conull, i.e. they have measure zero, or their complement has measure zero.

Lemma 8.4.5 *Let (x, y, z) be a point in $X_t(S_1)$ for $t > 2$ with some coordinate having magnitude less than 2. Then μ_t-almost every $(x', y', z') \in \Omega_t$ is equivalent to a point arbitrarily close to (x, y, z), in the Euclidean metric on $\mathbb{R}^3$.* ∎

PROOF Consider a Euclidean ball $D_\epsilon(x,y,z)$ of radius ϵ about (x,y,z) in $\mathbb{R}^3$. Take ϵ sufficiently small so that all points in $D_\epsilon(x,y,z)$ have a coordinate with magnitude less than 2. Consider the intersection

$$W_\epsilon(x,y,z) = B_\epsilon(x,y,z) \cap X_t(S_1).$$

By theorem 6.1.2, since $t > 2$, $X_t(S_1)$ is just the level set $\kappa^{-1}(t)$, and so $W_\epsilon(x,y,z)$ is a surface. The measure $\mu_t(W_\epsilon(x,y,z))$ is given by integrating the form ω, which can be written explicitly (see section 6.3) and is nowhere degenerate. Hence $\mu_t(W_\epsilon(x,y,z)) > 0$. Now the orbit $\Gamma \cdot W_\epsilon(x,y,z)$ cannot be null, hence by ergodicity is conull in Ω_t. This is true for arbitrarily small ϵ. Hence μ_t-almost every point in Ω_t is equivalent to a point in $W_\epsilon(x,y,z)$. ■

We obtain immediately the following more geometrically phrased result.

Lemma 8.4.6 *Let $\rho : \pi_1(S_1) \longrightarrow PSL_2\mathbb{R}$ be a representation taking some simple closed curve to an elliptic, and with character in Ω_t. Then μ_t-almost every representation ρ' with character in Ω_t has character equivalent to a character in $X_t(S_1)$ arbitrarily close to that of ρ.* ■

The idea is that ρ will be a fixed, "good" representation giving us the desired pentagon, hence cone-manifold structure, of proposition 8.4.2. It says that, from almost every ρ' with the stated properties, we may change basis and become arbitrarily close to ρ in character. The following result guarantees us "good" representations.

Lemma 8.4.7 *Let $t > 2$, $\delta > 0$ and a hyperbolic line $l \subset \mathbb{H}^2$, be given. Then there exists a representation $\rho_t^* : \pi_1(S_1) = \langle G, H\rangle \longrightarrow PSL_2\mathbb{R}$ and a point p within distance δ of l such that*

(i) $\mathrm{Tr}[g,h] = t$ *and* $\mathrm{Axis}[g^{-1},h^{-1}] = l$,

(ii) ρ_t^* *takes some simple closed curve (namely G) to an elliptic*

(iii) the pentagon $\mathcal{P}(g,h,p)$ is non-degenerate, bounds an embedded disc, and has a specified orientation.

PROOF Let $d = 2\cosh^{-1}(t/2)$; so d will be the translation distance of $[g,h] \sim [g^{-1},h^{-1}]$. Consider Fermi coordinates on l. Define the five points $p,q,r,s,t \in \mathbb{H}^2$ as follows:

$$p = (-d/2, \epsilon), \quad q = (-d/4, 0), \quad r = (0, -\epsilon), \quad s = (d/4, 0), \quad t = (d/2, \epsilon).$$

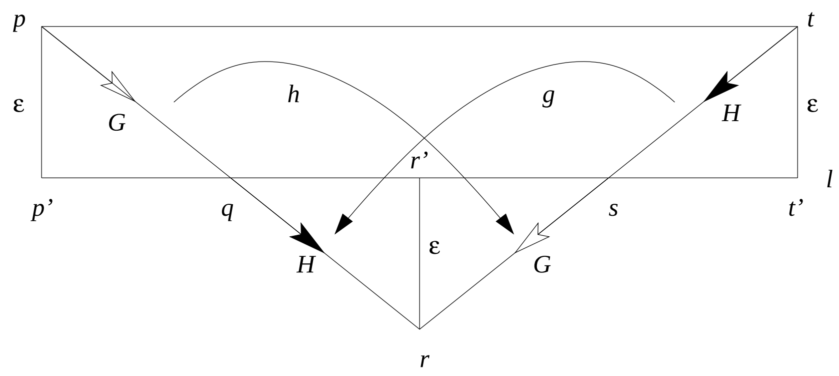

Figure 8.12: Construction of a good representation.

Let $p' = (-d/2, 0)$, $r' = (0, 0)$, $t' = (d/2, 0)$ respectively be the projections of p, r, t onto the line l, so the triangles $pp'q, rr'q, rr's, tt's$ are clearly congruent. It's clear that pqr and rst are geodesics. See figure 8.12.

Let g be the orientation-preserving isometry carrying the directed segment ts onto the segment qr. Let h be the orientation-preserving isometry carrying the directed segment pq onto the segment sr. Then

$$q = h^{-1}ghp, \quad r = ghp, \quad s = hp, \quad t = [g^{-1}, h^{-1}]p$$

so the pentagon $pqrst$ is actually the pentagon $\mathcal{P}(g, h, p)$, which is clearly non-degenerate and bounds an embedded disc. By replacing ϵ with $-\epsilon$ our pentagon could be of either possible orientation. This defines a representation $\rho_t^* : \langle G, H \rangle \longrightarrow PSL_2\mathbb{R}$. We claim that this ρ_t^* satisfies the required conditions.

As $\epsilon \longrightarrow 0$, the isometry g tends to a half-turn about the point $(d/8, 0)$, and h tends to a half-turn about the point $(-d/8, 0)$. Thus for ϵ sufficiently small both g, h are elliptics. We take ϵ which is sufficiently small and with $\epsilon < \delta$.

Take a unit tangent vector (p, u) at p in the direction of q. Chasing around unit vectors we see that $D(gh)(p, u)$ is a vector at $r = ghp$ also pointing along the line pqr. Thus gh translates along the line rp, from p to r. Similarly, $g^{-1}h^{-1}$ translates along the line rt, from r to t.

Now, q is the midpoint of pr and s is the midpoint of rt. Returning to the very first lemma of this thesis, lemma 3.2.1, we see that $[g^{-1}, h^{-1}] = (g^{-1}h^{-1})(gh)$ is a translation along the line qs, i.e. along l by twice the distance from q to s, i.e. by distance d. It follows that $|\operatorname{Tr}[g, h]| > 2$, but as g, h are elliptic, ρ_t^* cannot be the holonomy of a complete hyperbolic structure on S, which is discrete and maps entirely to hyperbolics. Hence $\operatorname{Tr}[g, h] > 2$ and $\operatorname{Tr}[g, h]/2 = \cosh(d/2) = t/2$. So $\operatorname{Tr}[g, h] = t$ as desired. ■

Note from the symmetry of our construction that g and h are conjugate via a reflection in a vertical axis of symmetry of figure 8.12. In particular, (choosing lifts of elements in $PSL_2\mathbb{R}$ to $SL_2\mathbb{R}$ appropriately) $\operatorname{Tr} g = \operatorname{Tr} h$, and so $(x, y, z) = (x, x, z)$.

Now we can complete the proof of proposition 8.4.2.

Proof (of 8.4.2) The idea: lemma 8.4.7 guarantees us a "standard good" representation ρ_t^*; for μ_t-almost any $\rho : \pi_1(S_1) \longrightarrow PSL_2\mathbb{R}$ with character in Ω_t we change basis so that the character approaches that of ρ_t^*, and conjugate ρ_t^* to another "good" ρ_0 which is actually close as a representation (not just as a character) to ρ.

So let $t > 2$ and $\epsilon > 0$, and an arbitrary basis G, H of $\pi_1(S_1)$ be given. Take an arbitrary line l in $\mathbb{H}^2$ and apply lemma 8.4.7 above to obtain a representation ρ_t^* and a point p^* within distance ϵ of l. Write $\rho_t^*(G) = g^*$, $\rho_t^*(H) = h^*$, and let the character of ρ_t^* be $(x^*, y^*, z^*) \in X_t(S_1)$. From our comment above $x^* = y^*$; from the lemma $|x^*| = |y^*| < 2$. And $\mathcal{P}(g^*, h^*, p^*)$ is non-degenerate, of the specified orientation, bounding an embedded disc.

Clearly ρ_t^* satisfies the hypotheses of lemma 8.4.6. Take a $\delta > 0$. Then by the lemma, μ_t-almost every ρ with character $(x, y, z) \in \Omega_t$ has $(x, y, z) \sim (x', y', z')$, where (x', y', z') is within δ of (x^*, y^*, z^*) in the Euclidean metric on $\mathbb{R}^3$.

Thus for μ_t-almost every ρ with character in Ω_t and every $\delta > 0$, we may choose a basis G', H' with $\rho(G') = g', \rho(H') = h'$ so that (x', y', z') is within δ of (x^*, y^*, z^*). If (G', H') does not have the same orientation as (G, H), replace (G', H') with (H', G'); the character changes as $(x', y', z') \mapsto (y', x', z')$ (since $\operatorname{Tr} gh = \operatorname{Tr} hg$). As $x^* = y^*$, the triple (y', x', z') is still within δ of (x^*, y^*, z^*). So we may take (G', H') to have the same orientation as (G, H).

Now the characters of ρ_t^* (with respect to the basis G, H) and ρ (with respect to the basis G', H') are $< \delta$ apart. Since $t > 2$, these characters correspond to unique conjugacy classes of representations (6.1.1). We may therefore conjugate ρ_t^* by some orientation-preserving isometry α to obtain a representation ρ_0 such that ρ_0 and ρ, with respect to appropriate bases, are close in $R(S_1)$ in some metric, say the Euclidean metric on $R(S_1) \subset \mathbb{R}^6$. We may even choose α so that $\operatorname{Axis}[g_0^{-1}, h_0^{-1}] = \operatorname{Axis}[g'^{-1}, h'^{-1}]$ and $[g_0^{-1}, h_0^{-1}]$ and $[g'^{-1}, h'^{-1}]$ translate in the same direction, thus are equal. We let $p_0 = \alpha(p^*)$ so p_0 lies within distance ϵ of $\operatorname{Axis}[g_0^{-1}, h_0^{-1}]$.

That is: for given $\eta > 0$, $\epsilon > 0$, $t > 2$, a basis G, H of $\pi_1(S)$, and for μ_t-almost every ρ with character in Ω_t, there exists a basis G', H' of $\pi_1(S)$, a representation ρ_0 of $\pi_1(S)$ and a $p_0 \in \mathbb{H}^2$ with the following properties:

(i) ρ_0 (w.r.t. the basis G, H) and ρ (w.r.t. the basis G', H') are within η in $R(S_1)$;

(ii) $[g'^{-1}, h'^{-1}] = [g_0^{-1}, h_0^{-1}]$, where $g_0 = \rho_0(G)$, $h_0 = \rho_0(H)$, $g' = \rho(G')$, $h' = \rho(H')$;

(iii) p_0 lies within distance ϵ of $\text{Axis}[g_0^{-1}, h_0^{-1}] = \text{Axis}[g'^{-1}, h'^{-1}]$;

(iv) (G', H') has the same orientation as (G, H); and

(v) $\mathcal{P}(g_0, h_0, p_0)$ is non-degenerate bounding an embedded disc and has a specified orientation.

It follows that the pentagons $\mathcal{P}(g_0, h_0, p_0)$ and $\mathcal{P}(g', h', p_0)$ are arbitrarily close. Hence $\mathcal{P}(g', h', p_0)$ is non-degenerate, bounds an embedded disc, has p_0 within ϵ of $\text{Axis}[g'^{-1}, h'^{-1}]$, and has the specified orientation. This completes the proof. ■

8.4.4 Piecing together character varieties

It remains to see how the character varieties and associated measures decompose when we cut and paste our surfaces. As our closed surface S is cut along a curve D into a punctured torus S_1 and another surface W, we obtain natural maps between spaces

$$\begin{array}{ccc} D & \rightarrow & S_1 \\ \downarrow & & \downarrow \\ W & \rightarrow & S \end{array}$$

and hence between character varieties

$$\begin{array}{ccc} X(D) & \leftarrow & X(S_1) \\ \uparrow & & \uparrow \\ X(W) & \leftarrow & X(S \mid D). \end{array}$$

Here the pushout

$$X(S \mid D) = \left\{ ([\rho_1], [\rho_W]) \in X(S_1) \times X(W) \mid [\rho_1|_{\pi_1(D)}] = [\rho_W|_{\pi_1(D)}] \in X(D) \right\}$$

is *not* the same as $X(S)$; for instance, for holonomy representations ρ_1, ρ_2 with the same trace along D, there are many possible representations on S corresponding to twisting around the curve D. However there is a natural map $X(S) \longrightarrow X(S \mid D)$, which is surjective.

Away from singularities, which have measure zero, the map $X(S) \longrightarrow X(S_1)$ is a submersion, since it can be taken to be a polynomial map, indeed a coordinate map, from a $(6g-6)$-dimensional set to a 3-dimensional set. Recall the character variety is defined by taking traces of a fixed set of curves on the surface S. As in chapter 6, we take for S_1 a set of standard curves (G, H, GH) on S_1, where G, H is a basis of $\pi_1(S)$. We can take the chosen curves on S to contain the chosen curves on S_1 so that the map $X(S) \longrightarrow X(S_1)$ is just a coordinate projection. Let the coordinates on $X(S_1)$ be (x,y,z), and let the coordinates on $X(S)$ be $(x,y,z,w_1,\ldots,w_k)$. As μ_S is absolutely continuous with respect to Lebesgue measure, there exists a real function f (a Radon-Nikodym derivative) such that for any Lebesgue measurable $A \subset X(S)$, we have

$$\mu_S(A) = \int_{(x,y,z,w_1,\ldots,w_k)\in X(S)\subset \mathbb{R}^{k+3}} \chi_A \; f \; d\lambda(x,y,z,w_1,\ldots,w_k)$$

where $d\lambda$ denotes the $(6g-6)$-dimensional Euclidean area form in $X(S) \subset \mathbb{R}^{k+3}$ and χ_A denotes the characteristic function of the set A.

I claim that the symplectic 2-forms on $X(S)$ and $X(S_1)$ are related by the natural map $j : X(S) \longrightarrow X(S_1)$. As described in section 4.5, the tangent space to $X(S)$ at a point $[\rho_0]$ is $H^1(S;\mathcal{B})$, where $\mathcal{B}$ is the bundle of coefficients over S associated with the $\pi_1(S)$-module $\mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho}$. The tangent space to $X(S_1)$ is likewise $H^1(S_1;\mathcal{B}_1)$ where $\mathcal{B}_1$ is the bundle of coefficients over S_1 associated with the $\pi_1(S_1)$-module $\mathfrak{sl}_2\mathbb{R}_{\mathrm{Ad}\,\rho_1}$, where ρ_1 is the induced homomorphism on S_1. Note $\mathcal{B}_1 = \mathcal{B}|_{S_1}$. So the natural map $\iota : S_1 \hookrightarrow S$ induces $\iota^* : H^1(S;\mathcal{B}) \longrightarrow H^1(S_1;\mathcal{B}_1)$, and by naturality of cup product (see e.g. [33]) we obtain a commutative diagram

$$\begin{array}{ccc} T_{[\rho]}X(S) \times T_{[\rho]}X(S) \cong H^1(S;\mathcal{B}) \times H^1(S;\mathcal{B}) & & \\ \downarrow \iota^* \times \iota^* & \overset{\cup}{\searrow} & \\ T_{[\rho_1]}X(S_1) \times T_{[\rho_1]}X(S_1) \cong H^1(S_1;\mathcal{B}_1) \times H^1(S_1;\mathcal{B}_1) & \overset{\cup}{\to} & \mathbb{R}. \end{array}$$

Proof (of 8.4.3) Recall $B = \cup_{D\in\mathcal{U}} B_D$ and $B_D = \cup_t B_{D,t}$. For a separating curve $D \in \mathcal{U}$, splitting S into a punctured torus S_1 and a surface W, we defined $B_{D,t} \subset X(S)$ to be the set of all characters of representations ρ such that

(i) $\mathcal{E}(\rho)[S] = -\chi(S) - 1$;

(ii) ρ takes some simple closed curve on S_1 to an elliptic (hence $\mathcal{E}(\rho_1)[S_1] = 0$ and $\mathcal{E}(\rho_W)[W] = -\chi(W)$);

(iii) with respect to some dual tree $\mathcal{T}$, the induced representation ρ_1 is bad.

(iv) with respect to some basis G, H of $\pi_1(S_1)$, $\mathrm{Tr}[g,h] = t > 2$

We will first show $\mu_S(B_D) = 0$ for given $D \in \mathcal{U}$.

Under the natural map $j : X(S) \longrightarrow X(S_1)$, the image of B_D is a set of characters of representations of $\pi_1(S_1)$ which are bad. Let $j(B_D) = A \subset X(S_1)$. The image of each set $B_{D,t}$ lies in $X_t(S_1) \subset X(S_1)$; letting $j(B_{D,t}) = A_t$ and letting μ_t denote the measure on $X_t(S_1)$ we have by proposition 8.4.2 $\mu_t(A_t) = 0$. Note that $B_D \subseteq \left(A \times \mathbb{R}^k\right) \cap X(S)$.

Since j is just a coordinate projection, using coordinates as above we have

$$\begin{aligned}\mu_S(B_D) &= \int_{(x,y,z,w_1,\ldots,w_k)\in X(S)\subset\mathbb{R}^{k+3}} \chi_{B_D}\ f\ d\lambda(x,y,z,w_1,\ldots,w_k)\\ &\le \int_{(x,y,z,w_1,\ldots,w_k)\in X(S)} \chi_{\left(A\times\mathbb{R}^k\right)\cap X(S)}\ f\ d\lambda(x,y,z,w_1,\ldots,w_k)\\ &= \int_{(w_1,\ldots,w_k)} \left(\int_{(x,y,z)\in X(S_1)} \chi_A\ f(x,y,z,w_1,\ldots,w_k)\ d\lambda(x,y,z)\right) d\lambda(w_1,\ldots,w_k).\end{aligned}$$

Thus it is sufficient to show that for any given $(w_1,\ldots,w_k)$, the inner integral is zero. Now we introduce the variable $t=\kappa(x,y,z)=\operatorname{Tr}\rho(D)$. The map $(x,y,z)\mapsto\kappa(x,y,z)=t$ is polynomial, hence measurable, so we may disintegrate the measure $d\lambda(x,y,z)$ over t and obtain a family of measures on the level sets $X_t(S_1)$ (for details see e.g. [53]). However, on the level set $X_t(S_1)=\kappa^{-1}(t)$, we have the symplectic 2-form ω_t for $X(S_1)$. The form ω_t describes the density of the measure on $X_t(S_1)$. But we have seen above that by naturality of the cup product, ω_t is just the projection of ω under the natural map $X(S)\longrightarrow X(S_1)$. Hence the integral over each $X_t(S_1)$ is just some multiple of ω_t, say $m_t\omega_t$, where m_t is a constant depending only on t. Hence for given $(w_1,\ldots,w_k)$ we have

$$\int_{(x,y,z)\in X(S)} \chi_A\ f\ d\lambda(x,y,z) = \int_{t\in\mathbb{R}} m_t \left(\int_{(x,y,z)\in X_t(S_1)} \chi_{A_t}\ f\ \omega_t\right)\ d\nu(t)$$

Here ν is some measure on $\mathbb{R}$, namely the pushforward measure of $d\lambda$ under κ, obtained in disintegrating the measure on $X(S_1)$. But integrating the symplectic form ω_t gives the original measure μ_t on each $X_t(S_1)$. Hence as each $\mu_t(A_t)=0$ we obtain

$$\int_{(x,y,z)\in X(S)} \chi_A\ f\ d\lambda(x,y,z) = \int_{t\in\mathbb{R}} m_t\left(\int_{A_t} f\ \omega_t\right)\ d\nu(t) = \int_{t\in\mathbb{R}} m_t\cdot 0\ d\nu(t) = 0.$$

Thus $\mu_S(B_D)$ is at most an integral over $(w_1,\ldots,w_k)$ of 0, hence is 0. Now $\mathcal{U}$ certainly has cardinality no greater than the fundamental group of S, hence is countable. So the union $B=\cup_D B_D\subset X(S)$ is a countable union of sets of measure zero, hence has measure zero. ∎

Bibliography

[1] William Abikoff, The Real Analytic Theory of Teichmüller Space (1989).

[2] Iain R. Aitchison, Spin Structures on 3-manifolds (unpublished).

[3] Alan F. Beardon, The Geometry of Discrete Groups (1983).

[4] Robert L. Benedetto and William M. Goldman, The Topology of the Relative Character Varieties of a Quadruply-Punctured Sphere, Experimental Mathematics 8 (1999) 85–103.

[5] Oleg Bogopolski, Decompositions of Fundamental Groups of Closed Surfaces into Free Constructions, Geometriae Dedicata 94: 63–89, 2002.

[6] B. H. Bowditch, Markoff triples and quasifuchsian groups, Proc. London Math. Soc. 77 (1978) 697–736.

[7] B. H. Bowditch, A proof of McShane's identity via Markoff triples, Bull. London Math. Soc. 28 (1996) 73–78.

[8] Martin R. Bridson and André Haefliger, Metric Spaces of Non-Positive Curvature, Springer-Verlag (1999).

[9] Kenneth S. Brown, Cohomology of Groups (1982).

[10] Peter Buser, Geometry and Spectra of Compact Riemann Surfaces, Birkhäuser, Boston (1992).

[11] D. Cooper, M. Culler, H. Gillet, D. D. Long, P. B. Shalen, Plane curves associated to character varieties of 3-manifolds, Invent. Math. 118, 47–84 (1994).

[12] Daryl Cooper, Craig D. Hodgson, Steven P. Kerckhoff, Three-dimensional Orbifolds and Cone-Manifolds, MSJ Memoirs, Mathematical Society of Japan, Volume 5, 2000.

[13] John H. Conway, The Sensual (Quadratic) Form, Mathematical Association of America (1997).

[14] Marc Culler, Lifting Representations to Covering Groups, Advances in Mathematics 59, 64–70 (1986).

[15] Marc Culler and Peter B. Shalen, Varieties of group representations and splittings of 3-manifolds, Ann. of Math. 117 (1983), 109–146.

[16] Nathan M. Dunfield, Cyclic surgery, degrees of maps of character curves, and volume rigidity for hyperbolic manifolds, Invent. Math., 136 (1999) 623–657.

[17] S. Eilenberg, Homology of spaces with operators, I, Trans. Amer. Math. Soc. 61 (1947), 378–417; errata, 62 (1947), 548.

[18] David Eisenbud, Ulrich Hirsch and Walter Neumann, Travsverse fliations of Seifert bundles and self homeomorphism of the circle, Comment. Math. Helvetici 56 (1981) 638–660.

[19] Y. Eliashberg and N. Mishachev, Introduction to the h-Principle, Graduate Studies in Mathematics Volume 48, American Mathematical Society, Providence, Rhode Island (2000)

[20] Stefano Francaviglia, Hyperbolicity equations for cusped 3-manifolds and volume-rigidity of representataions, PhD thesis, Scuola Normale Superiore Pisa 2003.

[21] R. Fricke, Über die Theorie der automorphen Modulgrupper, Nachr. Akad. Wiss. Göttingen (1896) 91–101.

[22] R. Fricke and F. Klein, Vorlesungen über die Theorie der automorphen Funktionen, vol. 1. Stuttgart, B.G. Teubner (1897).

[23] Daniel Gallo, Michael Kapovich, and Albert Marden, The monodromy groups of Schwarzian equations on closed Riemann surfaces, Annals of Mathematics, 151 (2000), 625–704.

[24] J. Gilman and B. Maskit, An Algorithm for 2-Generator Fuchsian Groups, Michigan Math. J. 38 (1991) 19–32.

[25] William M. Goldman, Discontinuous groups and the Euler class, PhD thesis, University of California, Berkeley (1980).

[26] William M. Goldman, The Symplectic Nature of Fundamental Groups of Surfaces, Advances in Mathematics 54, 200–225 (1984).

[27] William M. Goldman, Invariant functions on Lie groups and Hamiltonian flows of surface group representations, Invent. math. 85, 263–302 (1986).

[28] William M Goldman, Topological components of spaces of representations, Invent. math. 93, 557–607 (1988).

[29] William M Goldman, Ergodic theory on moduli spaces, Ann. Math. 146 (1997) 1–33.

[30] William M Goldman, The modular group action on real $SL(2)$-characters of a one-holed torus, Geometry & Topology, Volume 7 (2003) 443–486.

[31] Mikhael Gromov, Partial differential relations, Springer-Verlag (1980).

[32] P.J. Hilton and U. Stammbach, A Course in Homological Algebra (1971).

[33] C. D. Hodgson, Degeneration and regeneration of geometric structures on three-manifolds, PhD Thesis, Princeton University, 1986.

[34] Craig D. Hodgson and Steven P. Kerckhoff, Harmonic deformations of hyperbolic 3-manifolds, Kleinian Groups and Hyperbolic 3-Manifolds, Lond. Math. Soc. Lec. Notes 299, 41–73.

[35] Robert D. Horowitz, Induced automorphisms on Fricke characters of free groups, Trans. Amer. Math. Soc. 208 (1975), 41–50.

[36] J. Hoste and P. D. Shanahan, Trace Fields of Twist Knots, J. of Knot Theory and its Ramifications, 10 (2001), 625–639.

[37] J. Huebschmann, Symplectic and Poisson structures of certain moduli spaces I, Duke Math. J. 80 (1995), 737–756.

[38] Troels Jorgensen, On pairs of once-punctured tori, Kleinian Groups and Hyperbolic 3-manifolds, Lond. Math. Soc. Lec. Notes 299, 183–207 (2003).

[39] Xavier Leleu, Géométries de courbure constante des 3-variétés et variétés de caractères de représentations dans $SL_2(\mathbb{C})$, PhD Thesis, CMI, Université de Provence, Marseille, April 2000.

[40] R. Lyndon and R. Schupp, Combinatorial Group Theory, Springer–Verlag, Berlin, Heidelberg, New York (1977).

[41] Colin Maclachlan and Alan W. Reid, The Arithmetic of Hyperbolic 3-Manifolds, Springer (2000).

[42] W. Magnus, A. Karrass, D. Solitar, Combinatorial Group Theory: Presentations of groups in terms of generators and relations, Dover Publications, New York (1970).

[43] W. Magnus, Rings of Fricke characters and automorphism groups of free groups, Math. Z. 170, 91–103 (1980).

[44] W. Magnus, The uses of 2 by 2 matrices in combinatorial group theory. A survey. Result. Math. 4, 171–192 (1981).

[45] A. A. Markoff, Sur les formes binaires indefinies, I, Math. Ann. 15 (1879) 281–309; II, 17 (1880), 379–400.

[46] J. Peter Matelski, The classification of discrete 2-generator subgroups of $PSL(2,\mathbb{R})$, Israel J. Math. 42 (1982), 309–317.

[47] Daniel V. Mathews, Mahler's Unfinished Symphony: Etudes in Knots, Algebra and Geomtry, honours project, University of Melbourne 2003.

[48] Daniel V. Mathews, Quadratic Geography, Algebraic Extreme Sports, and Magical Farey Trees, Austral. Math. Soc. Gaz. 31 (2004) 147–156.

[49] John Milnor, On the existence of a connection with curvature zero, Comment. Math. Helv. 32 (1958) 215–223.

[50] John W. Milnor and James D. Stasheff, Characteristic classes, Annals of mathematics studies No. 76, Princeton university press (1974).

[51] J. Nielsen, Die isomorphismen der allgemeinen unendlichen Gruppe mit zwei Erzeugenden, Math. Ann. 71 (1918) 385–397.

[52] J. Nielsen, Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen I, Acta Math 50 (1927) 189–358.

[53] David Pollard, A User's Guide to Measure Theoretic Probability (2002).

[54] Norman Purzitsky, Two-Generator Discrete Free Products, Math. Z. 126, 209–223 (1972).

[55] Caroline Series, The Geometry of Markoff Numbers, The Mathematical Intelligencer, 7 (1985) 20–29.

[56] Norman Steenrod, The Topology of Fibre Bundles, Princeton University Press (1951).

[57] J. Stillwell, The Dehn-Nielsen Theorem, Appendix to "Papers on Group Theory and Topology by Max Dehn", Springer-Verlag, Berlin, Heidelberg, New York (1988).

[58] Ser Peow Tan, Branched $\mathbb{C}P^1$-structures on surfaces with prescribed real holonomy, Math. Ann. 300, 649–667 (1994).

[59] William P. Thurston, The geometry and topology of 3-manifolds, Mimeographed notes, Princeton University Mathematics Department 1979.

[60] William P. Thurston, Three-Dimensional Geometry and Topology: Volume 1, Edited by Silvio Levy, Princeton University Press, Princeton, New Jersey, 1997.

[61] Marc Troyanov, Prescribing curvature on compact surfaces with conical singularities, Transactions of the American Mathematical Society 324, 1991, 793–821.

[62] John W. Wood, Bundles with Totally Disconnected Structure Group, Comment. Math. Helv. 46 (1971), 257–273.