The Dynamics and Geometry of Kleinian Groups

Alexander Elzenaar

June 21, 2021

Contents

Introductory remarks

These notes are an introduction to the actions of Möbius transformations on hyperbolic space, and the manifolds which are constructed via taking quotients of the space by the group action.

Primary references are [6, 34]. See also [38] for geometric fundamentals.

Chapter 1

Möbius transformations

The goal for this chapter is to introduce the notion of a Möbius transformation on $\hat{\mathbb{R}}^n$ (which we shall, in the case $n = 2$, call a 'fractional linear transformation' and consider with the natural action on $\hat{\mathbb{C}}$).

1.1 Differentiability

Suppose $U \subseteq \mathbb{R}^n$ is open, and let $f : U \to \mathbb{R}^n$ be a function. We say that f is **differentiable** at some $x \in U$ if there exists a linear map $df_x : \mathbb{R}^n \to \mathbb{R}^n$ such that there exists an open neighbourhood V of x such that for all $y \in V$,

$$f(y) - f(x) = df_x(y - x) + o(y - x)$$

where o is a function with $o(0) = 0$ and $\lim_{y \to x} \frac{\|o(y-x)\|}{\|y-x\|} \to 0$.

We shall be interested in maps $\mathbb{R}^n \to \mathbb{R}^n$ and $\mathbb{C} \to \mathbb{C}$ which 'locally preserve angles'. We will wish to define this as 'have differentials which preserve angles'.

Let $A \in \mathrm{GL}(n, \mathbb{R})$. We say A is **orthogonal** if $(Ax, Ay) = (x, y)$ for all $x, y \in \mathbb{R}^n$. It is easy to prove (using the fact that A and A^t are adjoint operators) that this is equivalent to the condition $A^t A = I$, and that the set of all such operators forms a subgroup $O(n) \leq \mathrm{GL}(n, \mathbb{R})$. An orthogonal map is precisely a linear Euclidean isometry; and all Euclidean isometries are of the form $Ax + a$ for some $A \in O(n)$ and some $a \in \mathbb{R}^n$.

In the complex case, we say $A \in \mathrm{GL}(n, \mathbb{C})$ is **unitary** if $(Ax, Ay) = (x, y)$ for all $x, y \in \mathbb{C}^n$. Again, it is easy to prove that this is equivalent to the condition $A^* A = I$ (where A^* is the conjugate transpose of A), and that we have defined a subgroup $U(n) \leq \mathrm{GL}(n, \mathbb{C})$.

We say that f is **conformal** at x if the operator df_x is a non-zero scalar multiple of some orthogonal matrix; we say that f is **orientation preserving** or **reversing** according to whether $\det df_x$ is positive or negative.

We may identify $\mathbb{C}$ with $\mathbb{R}^2$ (in this case, differentiable maps are called **holomorphic**), and ask for $z \in U$ that the linear map df_z is in fact multiplication by some $w \in \mathbb{C}$; doing this, we deduce the **Cauchy-Riemann equations**,

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

(where we are writing $z = x + iy$ and $w = u + iv$, so $f(x + iy) = u + iv$). In any case, we see that

df_x has matrix representation

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

for $a = \frac{\partial u}{\partial x}$ and $b = \frac{\partial u}{\partial y}$. A simple computation then shows that $A^t A = (\det A)I$, so df_x is an orthogonal transformation whenever $\det df_x \neq 0$. In other words, a holomorphic map f is conformal away from its critical points. Note also, $\det A = a^2 + b^2 \geq 0$; so holomorphic maps are orientation preserving.

We briefly discuss now a central example of conformal maps.

1.1.1 Example. A **linear fractional transformation** is a map $f : \mathbb{C} \to \mathbb{C}$ such that

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

for all $z \in \mathbb{C}$, where $a, b, c, d \in \mathbb{C}$. Note that f is defined and holomorphic at every point except $z = -d/c$.

Consider now $\mathbb{PC}^1$ with homogenous coordinates; there is a natural action of $\mathrm{End}(2, \mathbb{C})$ on this space, namely by direct multiplication. Consider the computation

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} z \\ 1 \end{bmatrix} = \begin{bmatrix} az+b \\ cz+d \end{bmatrix}$$

which shows that this natural action gives us (whenever $z \neq -d/c$) precisely the action of f on the copy of $\mathbb{C}$ in $\mathbb{PC}^1$ with second coordinate 1. In this way we have a natural identification between the group of non-singular fractional linear transformations (those with $\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$) and the group $\mathrm{PSL}(2, \mathbb{C})$. It is natural then to view fractional linear transformations as acting on the Riemann sphere $\hat{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$, sending $-c/d \mapsto \infty$ and $\infty \mapsto a/c$ (with $a/c := \infty$ when $c = 0$). We will use the symbol $\mathbb{M}$ to denote the group of fractional linear transformations with their action on the Riemann sphere.

Now suppose f is singular, so $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = 0$. This implies that the kernel of the matrix A as it acts as a linear map on $\mathbb{R}^2$ is nontrivial. If $\dim \ker A = 2$, then A is the zero transformation and thus does not even act on $\mathbb{PC}^1$. On the other hand, if $\dim \ker A = 1$ then $\dim \operatorname{im} a = 1$; that is, f is a partial function on $\mathbb{PC}^1$ which sends all the points in its domain of definition to a single line in $\mathbb{R}^2$ and thus a single point in $\mathbb{PC}^1$.

Finally, we note that given any three points $z_1, z_2, z_3 \in \mathbb{C}$ and any other three points $w_1, w_2, w_3 \in \mathbb{C}$ there is a fractional linear transformation sending each z_i to the respective w_i (fractional linear transformations are **triply transitive**). Indeed, it suffices to show this when $(w_1, w_2, w_3) = (0, 1, \infty)$; and then the map

$$z \mapsto \frac{z - z_1}{z - z_3} \frac{z_2 - z_3}{z_2 - z_1}$$

works.

1.2 The general Möbius group

In this section, we follow essentially [6, section 3.1] and [14, chapter 5].

Let $a \in \mathbb{R}^n$ and $r > 0$. We write $S(a, r)$ for the sphere of radius r, centred at a. There is a natural notion of a reflection through $S(a, r)$.

1.2.1 Proposition. *For any $x \in \mathbb{R}^n \setminus \{a\}$, there is a unique point x' on the ray $\overrightarrow{ax}$ such that $\|x' - a\| \|x - a\| = r^2$.*

Proof. There is a unique intersection point between the ray and the sphere $S(a, r^2/\|x - a\|)$. □

We usually consider the sphere $S(a,r)$ as a subset of the space $\hat{\mathbb{R}}^n := \mathbb{R}^n \cup \{\infty\}$. In this case, we define a transformation $\phi : \hat{\mathbb{R}}^n \to \hat{\mathbb{R}}^n$ that sends each $x \in \mathbb{R}^n \setminus \{a\}$ to the point x' of $\mathbb{R}^n$ defined by Proposition 1.2.1, and which swaps a and ∞. We call this transformation the **sphere inversion** with respect to $S(a,r)$.

It is fairly easy to write down a formula for the action of ϕ on $\mathbb{R}^n \setminus \{a\}$:

1.2.2 Proposition. *If ϕ is sphere inversion with respect to $S(a,r)$, then*

$$\phi(x) := a + \left(\frac{r}{\|x - a\|}\right)^2 (x - a).$$

Proof.

$$\|\phi(x) - a\| = \left\| \left(\frac{r}{\|x - a\|}\right)^2 (x - a) \right\| = \frac{r^2}{\|x - a\|}.$$ □

It is clear by looking at this formula that ϕ is continuous at all points of $\hat{\mathbb{R}}^n$ except possibly a and ∞. We may place a topology on $\hat{\mathbb{R}}^n$ to make ϕ continuous at every point; in order to do this, we will need some geometric information about ϕ.

1.2.3 Lemma. *Let ϕ be a sphere inversion with respect to a sphere S; then ϕ is an involution of $\hat{\mathbb{R}}^n$ which exchanges the interior and the exterior of S.* □

Notation. If $x, y \in \mathbb{R}^n$, we use the notation $|x,y|$ for the Euclidean distance $\|x - y\|$. We shall use $d(\cdot,\cdot)$ to denote another metric on $\mathbb{R}^n$ (the **chordal metric**) later on.

1.2.4 Lemma. *Let $x, y \in \mathbb{R}^n$ be points, and let ϕ be sphere inversion with respect to $S(a,r)$; then*

$$|\phi(x), \phi(y)| = \frac{r^2}{|x,a||a,y|} |x,y|.$$

Proof. It is easy to see that the triangles axy and $a\phi(y)\phi(x)$ are similar (indeed, the angle at a is shared and one other angle on both triangles is a right angle). Then

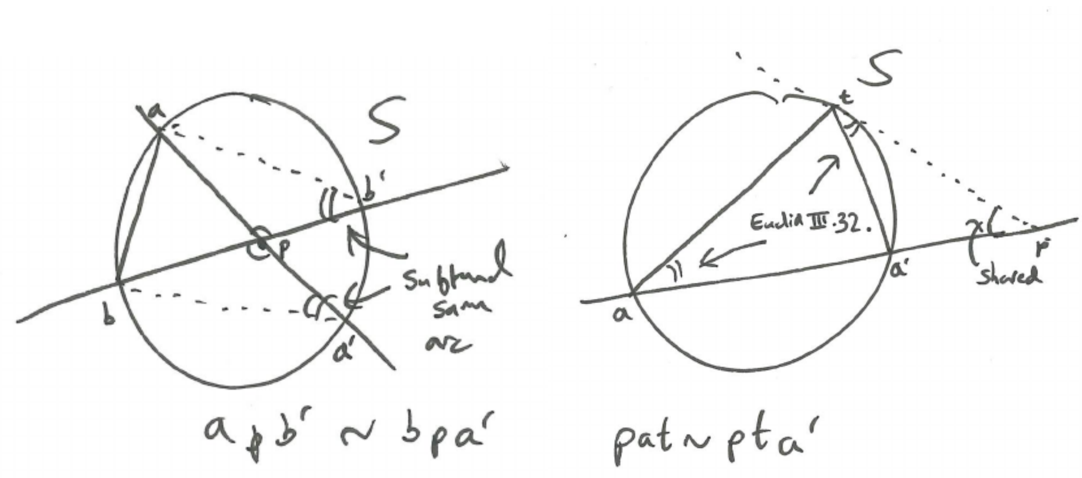
$$\frac{|\phi(x), \phi(y)|}{|x,y|} = \frac{|a, \phi(x)|}{|a,y|} = \frac{|a,x|}{|a,x|} \frac{|a,\phi(x)|}{|a,y|} = \frac{r^2}{|x,a||a,y|}$$

completes the proof. □

1.2.5 Theorem. *Let $H \subseteq \mathbb{R}^n$ be a sphere or a hyperplane (when we embed hyperplanes of $\mathbb{R}^n$ into $\hat{\mathbb{R}}^n$, we will always consider them as spheres through the point ∞), and let ϕ be the sphere inversion with respect to $S = S(a,r)$. Then $\phi(H)$ is a sphere or a hyperplane.*

More precisely:

1. *The image under ϕ of a hyperplane containing a is a hyperplane containing a.*
2. *The image under ϕ of a hyperplane not containing a is a sphere containing a.*
3. *The image under ϕ of a sphere containing a is a hyperplane not containing a.*



(a) Case that p lies inside S.

(b) Case that p lies outside S.

Figure 1.1: Proof of Lemma 1.2.6.

4. *The image under ϕ of a sphere not containing a is a sphere not containing a.*

We use the following lemma.

1.2.6 Lemma (Jakob Steiner, 1826)**.** *Let $S = S(a,r)$ be a sphere. If two lines through a point p meet S at points a, a' and b, b' respectively then $|p,a||p,a'| = |p,b||p,b'|$.*

We call the constant value of the product the **power** of p with respect to S.

Proof. If p is inside the sphere (Fig. 1.1a), we have similar triangles apb' and bpa' whence we obtain $|p,a|/|p,b'| = |p,b|/|p,a'|$. If p is outside the sphere (Fig. 1.1b), consider a point t such that the line pt is tangent to S. The triangles pta and $a'tp$ are similar, whence $|p,t|/|t,a| = |a',t|/|t,p|$; thus $|t,p|^2 = |t,a||t,a'|$, which is independent of the choice of a.

Proof of Theorem 1.2.5.

1. Let H be a hyperplane containing a; then for any $x \in H$, the ray $\overrightarrow{ax}$ lies in H; hence $\phi(x) \in H$. Since ϕ is its own inverse, $\phi^{-1}(x) \in H$ for all $x \in H$, and thus $\phi(\phi^{-1}(x)) = x$ exhibits x as the image of a point in H, i.e. $\phi(H)$ is bijective on H.

2. Let H be a hyperplane disjoint from a, let x be the foot of the perpendicular from H to a, and write x' for $\phi(x)$ (Fig. 1.2). We claim that $\phi(H)$ is the sphere S' with diameter $[a,x']$. Indeed, pick $y \in H$, and let y' be the intersection of the ray $\overrightarrow{ay}$ and the sphere S'. The triangles axy and $ay'x'$ are similar, hence
$$\frac{|a,y|}{|a,x|} = \frac{|a,x'|}{|a,y'|} \implies |a,y||a,y'| = |a,x'||a,x| = r^2$$
so $y' = \phi(y)$; as in (1) it is easy to see that ϕ is bijective between H and S'.

3. Apply (2) to ϕ^{-1}.

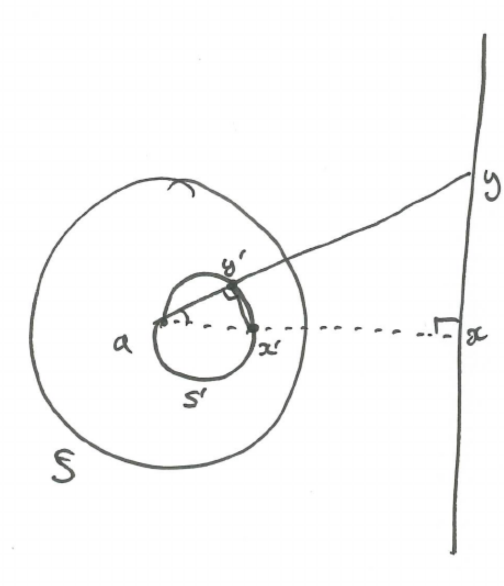


Figure 1.2: Image under inversion of a hyperplane not through the centre of inversion.

4. Surprisingly difficult; for a different approach, see [14, section 5.4] (via cross ratios). Let $S' = S(a', r')$ be a sphere not containing a; write p for the power of a with respect to S' (Lemma 1.2.6), and let ψ be the map $x \mapsto \frac{r^2}{p}(x-a)+a$. This is a Euclidean dilation with centre a, so the image $\psi(S')$ is a sphere S''. In particular, if x is any point of S' then $[\psi(a'), \psi(x)]$ is a radius of S''. We have by the properties of dilations that

$$\frac{|a, \psi(x)|}{|a, \psi(a')|} = \frac{r^2}{p}.$$

Let y be the other intersection point of $\overrightarrow{a, x}$ and S'. By the properties of the power of a, we have $|a, y||a, x| = p$. Substituting in the above display and cancelling, we have

$$\frac{|a, \psi(x)|}{|a, x|} = \frac{r^2}{|a, y||a, x|} \implies |a, \psi(x)||a, y| = r^2$$

and thus $\psi(x)$ is the image of x under the spherical transformation with respect to S. We have therefore shown that $\phi\restriction_{S'} = \psi$, and in particular $\phi(S') = S''$.

Based on this theorem, we will view hyperplanes as 'spheres through infinity' and hyperplane reflections as generalised sphere inversions. In particular, a **reflection** will be a map $\phi : \hat{\mathbb{R}}^n \to \hat{\mathbb{R}}^n$ which is either a sphere inversion or the reflection across a hyperplane (extended to fix ∞).

1.2.7 Proposition. *Let $\mathscr{B}$ be the set of subsets of $\hat{\mathbb{R}}^n$ consisting of the following:*

- *For each $r > 0$ and each $a \in \mathbb{R}^n$, the set $B(a, r) := \{x \in \mathbb{R}^n : ||x - a|| < r\}$;*
- *For each $r > 0$, the set $B(\infty, r) := \{x \in \mathbb{R}^n : ||x|| > r\} \cup \{\infty\}$.*

Then $\mathscr{B}$ is a basis for a topology on $\hat{\mathbb{R}}^n$.

Proof. Recall that a basis satisfies two properties: (1), for all $x \in \hat{\mathbb{R}}^n$, there exists $B \in \mathscr{B}$ such that $x \in B$; and (2), if $x \in B_1 \cap B_2$ for $B_1, B_2 \in \mathscr{B}$ then there exists $C \in \mathscr{B}$ such that $x \in C \subseteq B_1 \cap B_2$. Clearly (1) is satisfied. For (2), let $B_1, B_2 \in \mathscr{B}$; we have three possibilities. If B_1 and B_2 are Euclidean balls, then (2) is satisfied since Euclidean balls form a basis for the usual topology on $\mathbb{R}^n$. If $B_1 = B(\infty, r_1)$ and $B_2 = B(\infty, r_2)$ are balls about infinity, then $B_1 \cap B_2 = B(\infty, \max\{r_1, r_2\}) \in \mathscr{B}$. Finally, suppose $B_1 = B(a, r_1)$ and $B_2 = B(\infty, r_2)$ for $a \in \mathbb{R}^n$. Both of these have open intersection with $\mathbb{R}^n$, so $B_1 \cap B_2$ contains a Euclidean ball around each of its elements. ∎

We call the topology generated by $\mathscr{B}$ the **inversive topology** on $\hat{\mathbb{R}}^n$. Clearly $\mathbb{R}^n$ has the Euclidean topology as a subspace of $\hat{\mathbb{R}}^n$.

1.2.8 Proposition. *Let $\phi : \hat{\mathbb{R}}^n \to \hat{\mathbb{R}}^n$ be a reflection. Then ϕ is continuous with respect to the inversive topology.*

Proof. . It suffices to check that $\phi^{-1}(B) = \phi(B)$ is open for each $B \in \mathscr{B}$. For reflections, this is an easy exercise. The case of spherical inversions is also easy to do using Theorem 1.2.5 and Lemma 1.2.3: $\phi^{-1} = \phi$ sends interiors and exteriors of spheres (Euclidean open balls and open balls around infinity, respectively) to interiors and exteriors of spheres or open half-planes. ∎

We now show that the inversive topology is in fact a *metric* topology. We will do this by pulling back the chordal metric of the n-sphere via stereographic projection.

Consider the natural injection $\tilde{\cdot} : \hat{\mathbb{R}}^n \to \hat{\mathbb{R}}^{n+1}$ defined by $(a_1, \dots, a_n)^\sim := (a_1, \dots, a_n, 0)$ on $\mathbb{R}^n$ and sending $\infty \to \infty$. Let S^n denote the n-sphere embedded in $\hat{\mathbb{R}}^{n+1}$, namely $S^n = S(0,1) \subseteq \mathbb{R}^{n+1}$. Define the projection map $\pi : \mathbb{R}^{n+1} \to S^n$ by stereographic projection away from e^{n+1}; we find $\lambda \in (0,1]$ such that for $x \in \mathbb{R}^n$,

$$1 = \|\lambda\tilde{x} + (1-\lambda)e_{n+1}\|^2 = \lambda^2\|x\|^2 + (1-\lambda)^2 \implies 0 = \lambda^2\left(\|x\|^2 + 1\right) - 2\lambda$$

and thus $\lambda = \frac{2}{\|x\|^2+1}$; so

$$\pi(\tilde{x}) = \frac{2}{\|x\|^2+1}\tilde{x} + \frac{\|x\|^2-1}{\|x\|^2+1}e_{n+1}.$$

Defining $\pi(\infty) := \infty$, we obtain a bijection $\pi : \hat{\mathbb{R}}^n \to S^n$. We may define therefore a pullback metric on $\hat{\mathbb{R}}^n$, via

$$d(x,y) := \|\pi(\tilde{x}) - \pi(\tilde{y})\|. \tag{1.2.9}$$

This is the **chordal metric** on $\hat{\mathbb{R}}^n$.

1.2.10 Theorem. *The metric topology on $\hat{\mathbb{R}}^n$ induced by the chordal metric is precisely the inversive topology.*

Proof. We give an explicit formula for d:

$$d(x,y) = \begin{cases} \dfrac{2\|x-y\|}{(1+\|x\|)^2)^{1/2}(1+\|y\|)^2)^{1/2}} & x, y \in \mathbb{R}^n \\ \dfrac{2}{1+\|x\|^2} & y = \infty. \end{cases} \tag{1.2.11}$$

One can obtain this formula by realising that stereographic projection from $\mathbb{R}^n$ to S^n is precisely the sphere reflection in $S(e_{n+1}, \sqrt{2})$ and using the formulae $\|\tilde{x} - e_{n+1}\|^2 = 1 + \|x\|^2$ and Lemma 1.2.4.

It is easy to see now that every open set of one topology is open in the other. ∎

We now have a metric topology with the property that every reflection $\phi : \hat{\mathbb{R}}^n \to \hat{\mathbb{R}}^n$ is continuous with respect to it.

1.2.12 Definition. The group of homeomorphisms of $\hat{\mathbb{R}}^n$ generated by the set of reflections is the **general Möbius group**; an element of this group is a **Möbius transformation**. We use $\mathrm{GM}(n)$ to denote this group.

Since every Euclidean isometry is a finite composition of (hyperplane) reflections (this is proved geometrically in [13, chapter 3], and algebraically as [6, theorem 3.1.3]), we see that $\mathrm{Isom}(n) \leq \mathrm{GM}(n)$.

More interestingly, the fractional linear transformations of Example 1.1.1 are Möbius transformations.

1.2.13 Lemma. *The **complex inversion** map $z \mapsto z^{-1}$ is a Möbius transformation in $\hat{\mathbb{R}}^2$.*

Proof. Let $z = r\exp(i\theta)$; then $z^{-1} = r^{-1}\exp(-i\theta)$, so $z \mapsto z^{-1}$ is the composition of inversion in the unit circle and reflection across the real axis. $\square$

1.2.14 Proposition. *Let f be the fractional linear transformation given on $\hat{\mathbb{C}}$ by the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then*

$$f(z) = \frac{a}{c} - \frac{ad - bc}{c^2}\left(z + \frac{d}{c}\right)^{-1}$$

and so f is the composition of a translation, complex inversion, dilation, and a second translation. In particular, $f \in \mathrm{GM}(2)$. $\square$

We finally remark that there is a second natural generating set of $\mathrm{GM}(n)$, distinct from the set of all reflections.

1.2.15 Proposition. *The group $\mathrm{GM}(n)$ is generated by the set of transformations $\hat{\mathbb{R}}^n \to \hat{\mathbb{R}}^n$ consisting of:*

1. *The orthogonal transformations $O(n)$ (which are taken to fix ∞);*
2. *The sphere inversion with respect to $S(0, 1)$, $x \mapsto x^* := \|x\|^{-2}x$;*
3. *The real dilations about 0, $x \mapsto kx$ for all $k \in \mathbb{R}_{>0}$ (which are taken to fix ∞); and*
4. *The translations, $x \mapsto x + a$ for all $a \in \mathbb{R}^n$ (which are taken to fix ∞).*

Proof. It is clear (e.g. by the formula of Proposition 1.2.2) that a sphere inversion in $S(a, r)$ is precisely a composition of a translation $x \mapsto x - a$, sphere inversion through the unit sphere, dilation by r^2, and translation by a.

Note also, every Euclidean reflection is an orthogonal transformation composed with a translation. $\square$

1.3 Geometric properties of the general Möbius group

We already know that fractional linear transformations are conformal, being holomorphic. We now consider the general case. In this section, we follow essentially [6, section 3.2].

1.3.1 Theorem. *Every reflection is orientation-reversing and conformal. Thus every element of $\mathrm{GM}(n)$ is conformal; and an element is orientation-reversing or preserving as it is a product of an odd or an even number of reflections, respectively.*

Proof. Elementary but tedious computations with derivatives, see [6, theorem 3.1.6]. ∎

The following theorem essentially gives the converse to Proposition 1.2.14: every orientation-preserving Möbius transformation in $\hat{\mathbb{R}}^2$ is a fractional linear transformation.

1.3.2 Theorem. *Let $f : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ be a continuous injection, entire on $\mathbb{C}$. Then f is a fractional linear transformation.*

Proof. Let g be a fractional linear map sending $f(\infty) \mapsto \infty$, and let $h = gf$; then h is also a holomorphic injection of $\hat{\mathbb{C}}$. Since h fixes ∞, $h(\mathbb{C}) \subseteq \mathbb{C}$ and so $h\restriction_{\mathbb{C}}$ is an injective entire function. By Picard's little theorem ([2, theorem 5 of chapter 8]), since h is non-constant it has image $\mathbb{C}$ or $\mathbb{C} \setminus \{z\}$ for some $z \in \mathbb{C}$. On the other hand, h is a homeomorphism onto $h(\mathbb{C})$ and so the latter is simply-connected; this means $h(\mathbb{C}) = \mathbb{C}$ and h is bijective.

Now note, by continuity $\lim_{z\to\infty} h(z) = \infty$. Thus h has a pole at ∞, so $h(1/z)$ has a pole at 0. Thus we may expand $h(1/z)$ at 0 in the form $h(1/z) = z^{-k} \sum_{n=0}^{\infty} a_n z^n$ for some $k \in \mathbb{Z}$, and so $h(z) = z^k \sum_{n=0}^{\infty} a_n z^{-n}$ at 0. Since h does not have a pole at 0, there must be only positive powers in the expansion and so h is polynomial; and since h is injective, we must have that the degree of the polynomial is 1. Thus $h(z) = az + b$. Now $f = g^{-1}h$; and the composition of two fractional linear maps is fractional linear. ∎

1.3.3 Corollary. *Let $f : \hat{\mathbb{R}}^2 \to \hat{\mathbb{R}}^2$ be a Möbius transformation. Then:*

1. *If f is orientation preserving, then f is a fractional linear transformation.*

2. *If f is orientation reversing, then f is a **fractional reflection**: a map which acts on $\hat{\mathbb{C}}$ as*

$$z \mapsto \frac{a\bar{z} + b}{c\bar{z} + d}$$

for some $a, b, c, d \in \mathbb{C}$.

Proof. The orientation preserving case is precisely Theorem 1.3.2. Suppose f is orientation reversing; let g be the map $z \mapsto \bar{z}$, which is a reflection and hence Möbius. The composition fg is orientation preserving and so by part 1, $(fg)(z) = \frac{az+b}{cz+d}$ for $a, b, c, d \in \mathbb{C}$. Then $f(z) = (fgg)(z) = \frac{a\bar{z}+b}{c\bar{z}+d}$ as required. ∎

Recall, we defined $\mathbb{M}$ to be the group of fractional linear transformations acting on $\hat{\mathbb{C}}$. We will write $\mathrm{M}(n)$ for the subgroup of $\mathrm{GM}(n)$ of those Möbius transformations which are orientation preserving, with their natural action on $\hat{\mathbb{R}}^n$; identifying $\hat{\mathbb{C}}$ and $\hat{\mathbb{R}}^2$ gives a natural identification $\mathbb{M} = \mathrm{M}(2)$, and we will continue to avoid treating these group actions as being different.

1.3.4 Lemma. *Möbius transformations are transitive on spheres.*

Proof. Let $S \subseteq \hat{\mathbb{R}}^n$ be a sphere; it suffices to show that there is a transformation mapping S to the plane $x_n = 0$. If S is a plane, a Euclidean motion suffices. Otherwise, pick a point $x \in S$ and take a Möbius transformation sending $x \to \infty$, then this reduces to the plane case. ∎

1.3.5 Lemma. *Let σ be the Euclidean reflection in $\hat{\mathbb{R}}^n$ with respect to the plane S with equation $x_n = 0$. If ϕ is any Möbius transformation that fixes S pointwise, then $\phi = \sigma$ or ϕ is the identity.*

Proof. Let $a \in S$ and $r > 0$ and consider $S' = S(a,r)$; since $a, \infty \in S$, ϕ fixes a and ∞ and S' is mapped to a Euclidean sphere $S'' = S(b,s)$. Since S' is orthogonal to S, S'' is also orthogonal to S. Thus the centre of S'' must also lie on S. In particular, if x is a point of $S'' \cap S$ we have $|x,b| = s$; but $S'' \cap S = S' \cap S$ since ϕ fixes S pointwise, so such an x has $|x,a| = r$. Thus the points of $S'' \cap S = S' \cap S$ form a circle simultaneously of centre a and radius r, and centre b and radius s; this is nonsense unless $a = b$ and $r = s$. In particular, $S'' = S'$ and ϕ fixes S' (possibly not pointwise).

Let $x \in \hat{\mathbb{R}}^n \setminus S$, and let $y = \phi(x)$. Pick $a \in S$ and let S' be a sphere with centre a through x (so the radius r of S' is $||x - a||$). By the first paragraph, S' is fixed by ϕ, so $y \in S'$.

If $a = 0$, then we see that $||x|| = ||y||$. Taking $a = e_i$ for $i \in [n-1]$, we see that $||x - e_i|| = ||y - e_i||$ so $||x||^2 - 2(x, e_i) + 1 = ||y||^2 - 2(y, e_i) + 1$ for all i; cancelling, we see that $x_i = y_i$ for all such i. Further, $x_1^2 + \cdots + x_n^2 = y_1^2 + \cdots + y_n^2$ and so $x_n^2 = y_n^2$ and hence $x_n = \pm y_n$: so ϕ either fixes $\hat{\mathbb{R}}^n \setminus S$ pointwise or is the reflection across S. ⅏

1.3.6 Theorem. *Let σ the sphere inversion with respect to a sphere S. If ϕ is any Möbius transformation that fixes S pointwise, then $\phi = \sigma$ or ϕ is the identity.*

Proof. By Lemma 1.3.4 there is a Möbius transformation ψ sending S to the plane $x_n = 0$. Then the conjugation $\psi\sigma\psi^{-1}$ fixes the plane pointwise and is not the identity since σ is nontrivial. Hence by Lemma 1.3.5, $\psi\sigma\psi^{-1}$ is reflection across the plane.

Now note, by application of the same lemma we have that $\psi\phi\psi^{-1}$ is either the identity or reflection across the plane. If $\psi\phi\psi^{-1}$ is the identity, then ϕ is the identity. If $\psi\phi\psi^{-1} = \psi\sigma\psi^{-1}$, then $\phi = \sigma$. ⅏

Note, in the proof of the theorem we showed that any reflection σ is always conjugate via the map ψ given by Lemma 1.3.4 to the reflection across the plane $x_n = 0$. Thus

1.3.7 Corollary. *Any two reflections are conjugate in* GM(n). ⅏

1.3.8 Proposition. *Let $x, y, u, v \in \hat{\mathbb{R}}^4$ be distinct. The **cross-ratio** of these points is the number*

$$[x, y, u, v] = \frac{||x - u||||y - v||}{||x - y||||u - v||}.$$

A function $\phi : \hat{\mathbb{R}}^n \to \hat{\mathbb{R}}^n$ is a Möbius transformation iff it preserves cross ratios.

Proof. We use Proposition 1.2.15. It is clear that orthogonal maps and translations preserve the cross-ratio (they preserve distances), and that dilations preserve it (the scale factors cancel). Thus to show that all Möbius transformations fix the cross ratio it suffices to check the reflection ϕ in S^n. But note that by Lemma 1.2.4 we have

$$||\phi(x) - \phi(y)|| = \frac{||x - y||}{||x||||y||}$$

and the norms of the points x, y, u, v cancel symmetrically in the formula for the cross-ratio.

Suppose now that ϕ is a map preserving the cross-ratio. By composing ϕ with a suitable Möbius transformation we may assume that ϕ fixes ∞. For $x, y, u, v \in \mathbb{R}^n$, the ratio $[\infty, y, u, v]/[x, y, \infty, v]$ is invariant under ϕ. Hence

$$\frac{||\infty - u||||y - v||}{||\infty - y||||u - v||}\frac{||x - y||||\infty - v||}{||x - \infty||||y - v||} = \frac{||\infty - \phi(u)||||\phi(y) - \phi(v)||}{||\infty - \phi(y)||||\phi(u) - \phi(v)||}\frac{||\phi(x) - \phi(y)||||\infty - \phi(v)||}{||\phi(x) - \infty||||\phi(y) - \phi(v)||}$$

and cancelling we obtain

$$\frac{||\phi(x) - \phi(y)||}{x - y} = \frac{||\phi(u) - \phi(v)||}{u - v}$$

so ϕ is a Euclidean similarity and thus is a Möbius transformation. ⅏

1.3.9 Corollary. *If ϕ is a Möbius transformation which fixes ∞ then ϕ is a Euclidean similarity.* ⅏

1.4 Isometric spheres

Fix some $\phi \in \mathrm{GM}(n)$ which does not fix ∞; let $\alpha = \phi^{-1}(\infty)$ and let $\alpha' = \phi(\infty)$; hence ϕ maps the family $\mathcal{L}$ of spheres through α and ∞ to the family $\mathcal{L}'$ of spheres through ∞ and α'. Since ϕ is conformal, ϕ sends the family $\mathcal{S}$ of spheres orthogonal to the family $\mathcal{L}$ to the family $\mathcal{S}'$ of spheres orthogonal to $\mathcal{L}'$. Note, though, that these families are precisely the spheres around α and the spheres around α' respectively.

Note that ϕ acts continuously on the radii of the spheres of $\mathcal{S}$. Consider the natural action of ϕ on the closure of this set, $[0, \infty]$ (where 0 is the radius of the point-sphere around α and ∞ is the radius of the point-sphere around ∞). By the Brouwer fixed point theorem, the action on the set has a fixed point. However, ϕ swaps 0 and ∞. Thus there exists some $R \in [0, \infty)$ such that $\phi(S(\alpha, R)) = S(\alpha', R)$; and since ϕ is conformal, it preserves the chordal metric on $S(\alpha, R)$ (as the chordal metric between two points x and y is determined only by the radius R and the angle of the triangle $x\alpha y$.).

In fact, we may be more precise: we will show that the action of ϕ on $[0, \infty]$ is *monotone decreasing* and so the fixed point is unique. We will also compute explicitly the radius R.

1.4.1 Lemma. *Let $\phi \in \mathrm{GM}(n)$ fix 0 and leave the unit ball B^n invariant. Then $\phi \in O(n)$.*

Proof. Note by continuity that ϕ leaves S^{n-1} invariant; let σ be the reflection in S^n. Then $\phi^{-1}\sigma\phi$ fixes S^n and by Theorem 1.3.6 either $\phi^{-1}\sigma\phi = \sigma$ or $\phi^{-1}\sigma\phi$ is the identity. Note, $\phi^{-1}\sigma\phi(0) = \phi^{-1}\sigma(0) = \phi^{-1}(\infty)$; since $\phi(0) \neq \infty$, it therefore cannot be the case that $\phi^{-1}\sigma\phi(0) = 0$; so $\phi^{-1}\sigma\phi = \sigma$, and σ and ϕ commute. Thus $\phi(\infty) = \phi\sigma(0) = \sigma\phi(0) = \sigma(0) = \infty$; and by Corollary 1.3.9, ϕ is a Euclidean similarity. Since ϕ leaves S^{n-1} invariant and fixes the origin, it is immediate that $\phi \in O(n)$. ∎

1.4.2 Theorem. *Let $\phi \in \mathrm{GM}(n)$.*

1. *If ϕ fixes ∞, then there exists $A \in O(n)$, $r > 0$, and $x_0 \in \mathbb{R}^n$ such that $\phi(x) = r(Ax) + x_0$ for all $x \in \mathbb{R}^n$.*
2. *If ϕ does not fix ∞, then there exists $A \in O(n)$, $r > 0$, $x_0 \in \mathbb{R}^n$, and a sphere inversion σ such that $\phi(x) = rA\sigma x + x_0$ for all $x \in \mathbb{R}^n$.*

Proof. For (1), let $x_0 = \phi(0)$ and let r be the radius of $\phi(S^{n-1})$. Then $\psi : \hat{\mathbb{R}}^n \to \hat{\mathbb{R}}^n$ defined by $\psi(x) = (x - x_0)/r$ has the property that $\psi\phi$ fixes ∞ and preserves B^n. By Lemma 1.4.1, the map $\psi\phi \in O(n)$ and the result follows. For (2), compose ϕ with a sphere inversion σ sending $\infty \mapsto \phi^{-1}(\infty)$. ∎

1.4.3 Corollary. *Let $\phi \in \mathrm{GM}(n)$. There is a unique $\alpha \in \mathbb{R}^n$ and $R > 0$ such that ϕ acts isometrically on $S(\alpha, R)$.*

Proof. By the theorem, we may write $\phi(x) = rA\sigma x + x_0$ for some σ a reflection with respect to some sphere $S(\alpha, s)$. Note that α is necessarily $\phi^{-1}(\infty)$. By Lemma 1.2.4,

$$||\phi(x) - \phi(y)|| = r||\sigma(x) - \sigma(y)|| = \frac{rs^2||x - y||}{||x - \alpha||||\alpha - y||}.$$

Let $R = s\sqrt{r}$, and let $S = S(\alpha, R)$. Observe that the value

$$\lim_{y \to x} \left\| \frac{\phi(y) - \phi(x)}{y - x} \right\| = \frac{R^2}{||x - \alpha||||\alpha - y||}$$

is equal to 1 if and only if $x \in S$. Further, note that the value is greater than 1 precisely when x is inside S, and less than 1 when x is outside S; and this shows the monotonicity property mentioned above with respect to radii of points about α. ∎

1.5 The half-plane model of hyperbolic space

There is a natural action of $\mathrm{GM}(n)$ on the space $\hat{\mathbb{R}}^{n+1}$. We discuss this action now, essentially following [6, section 3.3].

1.5.1 Lemma. *Given a sphere $S \subseteq \hat{\mathbb{R}}^n$ and an affine embedding $j : \mathbb{R}^n \hookrightarrow \mathbb{R}^{n+1}$, there is a unique sphere S^j with the following properties:*

1. *S^j is orthogonal to $j(\mathbb{R}^n)$;*
2. *$S^j \cap j(\mathbb{R}^n) = S$.*

Proof. If $S = S(a, r)$ then define $S^j \coloneqq S(j(a), r)$; if S is a hyperplane through u with normal vector n then define S^j to be the hyperplane through $j(u)$ with normal vector $j(n)$. Uniqueness is evident: suppose S is a sphere, then S^j clearly must be a sphere and the only spheres orthogonal to an affine hyperplane have their centre on the hyperplane; if S is a hyperplane, for S^j to be orthogonal to $j(\mathbb{R}^n)$ the normal vectors of S^j must lie in $j(\mathbb{R}^n)$. ∎

1.5.2 Proposition. *Given an affine embedding $j : \mathbb{R}^n \hookrightarrow \mathbb{R}^{n+1}$ and a choice of a preferred half-space H^+ bounded by $j(\mathbb{R}^{n+1})$, for every $\phi \in \mathrm{GM}(n)$ there exists a unique $\phi^j \in \mathrm{GM}(n+1)$ preserving H^+ such that $\phi^j(js) = j(\phi^s)$ for all $s \in \hat{\mathbb{R}}^n$. This is the* ***Poincaré extension*** *of ϕ.*

Proof. Suppose first that ϕ is a reflection across some sphere $S \subseteq \hat{\mathbb{R}}^n$. Define ϕ^j to be the reflection across S^j in $\hat{\mathbb{R}}^{n+1}$. The ϕ^j clearly extend the action of ϕ in the following way: $\phi^j(js) = j(\phi(s))$ for all $s \in \hat{\mathbb{R}}^n$. It is also clear that the ϕ^j preserve the half-spaces of $j(\mathbb{R}^{n+1})$.

We may decompose an arbitrary $\phi \in \mathrm{GM}(n)$ into a product $\phi = \sigma_1 \cdots \sigma_k$ where each σ_i is a reflection; then $\phi^j \coloneqq \sigma_1^j \cdots \sigma_k^j$ clearly has the correct properties. Further, if ϕ_1^j and ϕ_2^j are two extensions of ϕ in the given sense then $(\phi_1^j)^{-1}\phi_2^j$ fixes each point of the hyperplane $j(\mathbb{R}^n)$ and thus by Theorem 1.3.6 we have $(\phi_1^j)^{-1}\phi_2^j$ is either reflection across $j(\mathbb{R}^n)$ or the identity; but $(\phi_1^j)^{-1}$ and ϕ_2^j preserve H^+ so their product does too, and so the product is the identity. ∎

Since $\cdot^j$ preserves composition, we therefore have a group embedding $\mathrm{GM}(n) \hookrightarrow \mathrm{GM}(n+1)$ for every embedding j. We now fix a preferred embedding j, and thus a preferred group embedding.

1.5.3 Definition. If $x \in \mathbb{R}^n$, define $\tilde{x}$ to be the point $(x_1, \dots, x_n, 0) \in \mathbb{R}^{n+1}$ and define $\tilde{\infty} = \infty$. Further, define H^{n+1} to be the half-space (excluding ∞)

$$\{(x_1, \dots, x_n, y) \in \mathbb{R}^{n+1} : y > 0\}$$

and let $\tilde{\phi}$ denote the image of $\phi \in \mathrm{GM}(n)$ under the group embedding induced by $\tilde{\cdot}$ and the choice of the preferred half-space H^{n+1}.

1.5.4 Lemma. *Let $x, y \in H^{n+1}$. Then, for all $\phi \in \mathrm{GM}(n)$, we have*

$$\frac{\|y - x\|^2}{y_{n+1}x_{n+1}} = \frac{\left\|\tilde{\phi}y - \tilde{\phi}x\right\|^2}{(\tilde{\phi}y)_{n+1}(\tilde{\phi}x)_{n+1}}.$$

Proof. It suffices to show that the equality holds whenever ϕ is a reflection. If ϕ is inversion in $S(a, r)$ then we have by Proposition 1.2.2 that

$$(\tilde{\phi}x)_{n+1} = \tilde{a}_{n+1} + \frac{r^2}{\|x - \tilde{a}\|^2}(x - \tilde{a})_{n+1} = \frac{r^2 x_{n+1}}{\|x\|^2};$$

from Lemma 1.2.4 we then have

$$\frac{\left\|\tilde{\phi}y-\tilde{\phi}x\right\|^2}{(\tilde{\phi}y)_{n+1}(\tilde{\phi}x)_{n+1}} = \frac{r^4\|x-y\|^2}{\|x-\tilde{a}\|^2\|\tilde{a}-y\|^2}\frac{1}{(\tilde{\phi}y)_{n+1}(\tilde{\phi}x)_{n+1}}$$
$$= \frac{r^4\|x-y\|^2}{\|x\|^2\|y\|^2}\frac{\|x\|^2\|y\|^2}{r^4x_{n+1}y_{n+1}}$$

which completes the proof. ∎

In fact, we can do even better:

1.5.5 Theorem. *Endow H^{n+1} with the Riemann metric g given by*

$$g := \frac{(dx_1)^2+\cdots+(dx_{n+1})^2}{x_{n+1}^2}.$$

(compare [32, theorem 3.7]). Then GM(n) *acts as a group of isometries of H^{n+1}, in the sense that $g=\phi^*g$ for all $\phi \in$* GM(n).

Proof. We neglect the proof, just use Lemma 1.2.4 to compute $d\phi_x$ and then compute the pullback in the same manner as Lemma 1.5.4. ∎

1.5.6 Definition. The space H^{n+1} endowed with the metric g of Theorem 1.5.5 is called the **half-plane model of hyperbolic $(n+1)$-space**. We write $\rho(x,y)$ for the distance between $x, y \in H^{n+1}$ with respect to g.

Further, we remark without proof that geodesics in H^{n+1} are of two kinds: Euclidean semicircles orthogonal to $\mathbb{R}^n$, and Euclidean rays orthogonal to $\mathbb{R}^n$.

Recall that the function $\cosh : \mathbb{C} \to \mathbb{C}$ is defined by

$$\cosh z := \frac{1}{2}\left(\exp(x)+\exp(-x)\right).$$

The following theorem gives a 'global' formula for the Riemannian metric in the half-space model.

1.5.7 Theorem. *If $x, y \in H^{n+1}$ then*

$$\cosh\rho(x,y) = 1+\frac{\|x-y\|^2}{2x_{n+1}y_{n+1}}.$$

Proof. Note first that the coordinates of the Riemann metric g are

$$g_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i=j\leq n \\ x_{n+1}^{-2} & i=j=n+1. \end{cases}$$

Hence $\det g = x_{n+1}^{-2}$, and the Riemann volume form is $dV_g = \sqrt{\det g}\,dx_1\wedge\cdots\wedge dx_{n+1} = \left|x_{n+1}\right|^{-1}dx_1\wedge\cdots\wedge dx_{n+1}$.

Suppose now that $x = se_{n+1}$ and $y = te_{n+1}$. The geodesic joining x and y is the Euclidean segment γ through them; thus

$$\rho(x,y) = \int_\gamma \left|x_{n+1}\right|^{-1}dx_1\wedge\cdots\wedge dx_{n+1} = \int_s^t \left|x_{n+1}\right|^{-1}dx_{n+1} = \left|\log t/s\right|,$$

and by direct substitution,

$$\cosh \rho(x,y) = \frac{1}{2}\left(\exp(\log t/s) + \exp(-\log t/s)\right) = \frac{1}{2}\left(\frac{t}{s} + \frac{s}{t}\right) = \frac{t^2+s^2}{2st} = 1+\frac{(t-s)^2}{2st} = 1+\frac{\|x-y\|^2}{2x_{n+1}y_{n+1}}.$$

Now, let y and y' be arbitrary. Let S be the (unique) Euclidean circle containing y and y' which is orthogonal to the hyperbolic line at infinity, $\hat{\mathbb{R}}^n$ (we continue to use the half-plane model, so this is the line $\{x_{n+1} = 0\}$). Let α and β be the intersection points $S \cap \hat{\mathbb{R}}^n$; at least one of these is not ∞, say $\alpha \neq \infty$. Let $\sigma \in \mathrm{GM}(n)$ be sphere inversion in $S(\alpha, 1)$; this sends $\alpha \mapsto \infty$, and sends $\beta \mapsto \sigma(\beta)$. In particular, the map $g \in \mathrm{GM}(n)$ given by $g(x) = \sigma(x) - \sigma(\beta)$ sends $(\alpha, \beta) \mapsto (\infty, 0)$, and the Poincaré extension $\tilde{g}$ must therefore send S onto some geodesic in H^3 with points at infinity $0, \infty \in \hat{\mathbb{R}}^n$. The only such geodesic is the x_{n+1}-axis, and so the map g sends y and y' onto the x_{n+1}-axis, say $g(y) = se_{n_1}$ and $g(y') = te_{n+1}$. Now note that both sides of the equality in the statement are preserved by the action of $\mathrm{GM}(n)$, the left side because $\mathrm{GM}(n)$ is a group of isometries and the right side by Lemma 1.5.4. In particular,

$$\begin{aligned}\cosh \rho(x,y) &= \cosh \rho(gx, gy) = \cosh \rho(se_{n+1,te_{n+1}}) \\ &= 1 + \frac{(t-s)^2}{2st} = 1 + \frac{\|gx - gy\|^2}{2(gx)_{n+1}(gy)_{n+1}} = 1 + \frac{\|x-y\|^2}{2x_{n+1}y_{n+1}}\end{aligned}$$

where we used our above computation that the result held on the x_{n+1}-axis. ♘⥄

1.6 The ball model of hyperbolic space

See also [6, section 3.4].

Let $\pi : \hat{\mathbb{R}}^n \to S^n$ be the usual stereographic projection from the north pole e_{n+1} of S^n, acting on the preferred embedding $\hat{\mathbb{R}}^{n^\sim}$. Recall that π acts as the restriction of the sphere inversion ϕ with respect to the sphere $S(e_{n+1}, \sqrt{2})$. Since $\phi(e_{n+1}) = \infty$ and ϕ maps $\hat{\mathbb{R}}^{n^\sim}$ onto S^n, we must have that the upper half-space is mapped to the exterior of S^n (the connected component of $\hat{\mathbb{R}}^n \setminus S^n$ containing ∞) and the lower half-space is mapped to the open ball B^{n+1}. Thus, if σ denotes reflection in $\hat{\mathbb{R}}^{n^\sim}$, the composition $f := \phi\sigma$ is a Möbius transformation mapping $H^{n+1} \xrightarrow{\sim} B^{n+1}$.

We may give an explicit formula for this composition:

$$f(x) = e_{n+1} + \frac{2}{\|\sigma x - e_{n+1}\|^2}(\sigma x - e_{n+1}) = e_{n+1} + \frac{2}{\|x - 2x_{n+1} - e_{n+1}\|^2}(x - 2x_{n+1} - e_{n+1}). \tag{1.6.1}$$

One can show that f induces the following Riemann metric on B^{n+1} from the hyperbolic Riemann metric on H^{n+1}:

$$g := 2\frac{(dx_1)^2 + \cdots + (dx_{n+1})^2}{1 - \|x\|^2}. \tag{1.6.2}$$

This is the **ball model of hyperbolic $n+1$-space**. The canonical picture is included as Fig. 1.3.

Since we define the metric on B^{n+1} in this way, every isometry of B^{n+1} is of the form fgf^{-1}, where g is an isometry of H^{n+1}. Because $f \in \mathrm{GM}(n+1)$, the group $\mathrm{GM}(n)$ is conjugate in $\mathrm{GM}(n+1)$ to the subgroup of $\mathrm{GM}(n+1)$ leaving B^{n+1} invariant.

Compare Eq. (1.6.2) with Eq. (1.2.11). We have the following philosophy:

- The preservation of the ball metric of H^{n+1} by $\mathrm{GM}(n)$ is equivalent to the preservation of the chordal metric of $\partial H^{n+1} = S^n = \hat{\mathbb{R}}^n$.

Figure 1.3: M. C. Escher, *Circle Limit I*, 1958.

- The preservation of the half-space metric of H^{n+1} is equivalent to the semi-preservation of the usual metric of ∂H^{n+1} via Lemma 1.2.4.

The general idea of the remainder is to study the action of special subgroups (**Kleinian groups**) $G \leq \mathbb{M}$ on sets of the form

$$H^3 \cup \Omega(G)$$

where $\Omega(G)$ is a 'nice' subset (the **regular set**) of $\partial H^2 = \hat{\mathbb{C}}$, and the quotient spaces $H^3 \cup \Omega(G)/G$ obtained by gluing together elements of the boundary of ∂H^2 according to this action. We shall see that the resulting space is very general.

Chapter 2

Classification of fractional linear transformations

In this chapter, we reduce to the case $n = 2$; thus we will consider only Möbius transformations which act on the Riemann sphere. We will follow a selection of topics from [34, chapter I] and [6, chapter 4].

2.1 Matrix analysis

See [6, section 2.2].

Recall that we may identify $\mathbb{M}$ and $\mathrm{PSL}(2, \mathbb{C})$. In order to study $\mathbb{M}$, we will need some invariants of elements of $\mathbb{M}$; the easiest way to define numerical invariants is in terms of the matrices that we may write down to represent them. To this end, we study the space of matrices $\mathrm{End}(2, \mathbb{C})$.

2.1.1 Lemma. *The map* $[A, B] \mapsto \mathrm{tr}(AB^*)$ *is an inner product on* $\mathrm{End}(2, \mathbb{C})$ *as a complex vector space.*

Proof. Note that $\mathrm{tr}(AB^*) = A_{11}\overline{B}_{11} + A_{12}\overline{B}_{12} + A_{21}\overline{B}_{21} + A_{22}\overline{B}_{22}$, so $\mathrm{tr}(AB^*)$ is just another way of writing the usual inner product on $\mathbb{C}^4 = \mathbb{C}^{2\times 2}$.

We therefore obtain a norm $\|A\| := \sqrt{[A, A]}$ and thus a metric topology on $\mathrm{End}(2, \mathbb{C})$. We have some additional relations beyond the usual axioms for a norm and metric.

2.1.2 Lemma. *Let* $A, B \in \mathrm{End}(2, \mathbb{C})$. *Then*

1. $|\det A| \left\|A^{-1}\right\| = \|A\|$,
2. $|[A, B]| \leq \|A\| \|B\|$,
3. $|AB| \leq |A||B|$, *and*
4. $2|\det A| \leq \|A\|^2$.

Proof. By direct computation with coordinates we obtain (1). Introduce the auxillary quantity $C = [B, A]A - \|A\|^2 B$; since $\|\cdot\|$ is a norm we have

$$\left[[B, A]A - \|A\|^2 B, [B, A]A - \|A\|^2 B\right] = \|C\|^2 \geq 0$$

and expanding the left hand size via the inner product axioms gives (2).

Suppose $AB = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$. Then $|p|^2 = |A_{11}B_{11} + A_{12}B_{21}|^2$; this latter is $\|(x,y)\|^2$ for $x = (A_{11}, A_{12})$ and $y = (B_{11}, B_{21})$ and so by Cauchy-Schwartz we havs $|p|^2 \le (|A_{11}|^2 + |A_{12}|^2)(|B_{11}|^2 + |B_{21}|^2)$. Writing similar formulae for $|q|^2, |r|^2, |s|^2$ allows us to write $\|AB\|^2 = |p|^2 + \cdots + |s|^2$ in terms of the entries of A and B; these entries in the sum factor to give $|A||B|$.

Finally for (4) note that $|\det A| \le |A_{11}A_{22}| + |A_{12}A_{21}|$ by the triangle inequality, so

$$\|A\|^2 - 2|\det A| \ge |A_{11}|^2 + \cdots + |A_{22}|^2 - 2(|A_{11}A_{22}| + |A_{12}A_{21}|) = (\|A_{11}\| - \|A_{22}\|)^2 + (\|A_{12}\| - \|A_{21}\|)^2 \ge 0.$$

This proves the lemma. ∎

We finally give some convergence results.

1. A sequence $(A_n : n \in \mathbb{N})$ in $\mathrm{End}(2,\mathbb{C})$ converges iff the sequences $((A_n)_{i,j} : n \in \mathbb{N})$ all converge.
2. The functions $\|\cdot\|$, tr, and det are all continuous with respect to the metric topology: $\|\cdot\|$ by definition, and tr and det because they are polynomial in the coefficients so can be viewed as polynomial maps $\mathbb{C}^4 \to \mathbb{C}$.
3. In $\mathrm{GL}(2,\mathbb{C})$, the map $A \mapsto A^{-1}$ is continuous (it is polynomial in the entries) and $A_nB_n \to AB$ for $(A_n), (B_n)$ sequences (because it converges componentwise); thus $\mathrm{GL}(2,\mathbb{C})$ is a topological group with respect to the metric.

2.2 The norm

Recall that we may identify $\mathbb{M}$ and $\mathrm{PSL}(2,\mathbb{C})$.

2.2.1 Lemma. *Let $g \in G$, and let $A, B \in \mathrm{SL}(2,\mathbb{C})$ be representatives of g in $\mathrm{PSL}(2,\mathbb{C})$. Then*

1. $\|A\| = \|B\|$;
2. $\mathrm{tr}^2 A = \mathrm{tr}^2 B$.

Of course, $\mathrm{tr}\, A$ is not well-defined!

Proof. Note, for $A \in \mathrm{GL}(2,\mathbb{C})$ we have $\det(\lambda A) = \lambda^2 A$. If A and B are equivalent matrices in $\mathrm{PSL}(2,\mathbb{C})$ then $\lambda A = B$; but $\det A = 1 = \det B$, so $\lambda^2 = 1$ and $\lambda = \pm 1$. Thus either $A = B$ (in which case $\|A\| = \|B\|$ and $\mathrm{tr}^2 A = \mathrm{tr}^2 B$ trivially) or $A = -B$ (in which case $\|A\| = \|B\|$ since the norm depends on absolute values of components only, and $\mathrm{tr}^2 A = \mathrm{tr}^2 B$ since $(A_{11} + A_{22})^2 = (-A_{11} - A_{22})^2$). ∎

In this section, we study the first of these invariants (the norm). It turns out that the trace squared is a more useful invariant, as it will allow us to detect global geometric properties of the group elements. We shall study it in the next section.

Recall that GM(2) and hence $\mathbb{M}$ acts naturally on $\hat{\mathbb{R}}^3$ and H^3. It will be convenient to describe the geometry of this action in terms of the quaternions.

2.2.2 Definition. The **quaternion algebra** is the algebra $\mathbb{H}$ generated by the set $\{\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ over $\mathbb{R}$ with the relations $i^2 = j^2 = k^2 = ijk = -1$. The congugate $\overline{q}$ of $q = x\mathbf{1} + y\mathbf{i} + u\mathbf{j} + v\mathbf{k}$ is defined to be $x\mathbf{1} - y\mathbf{i} - u\mathbf{j} - v\mathbf{k}$. We identify the complex numbers with the image of the injection $x + yi \mapsto x\mathbf{1} + y\mathbf{i}$.

Note, if $z = x + yi$ and $w = u + vi$ then $z + w\mathbf{j} = x + y\mathbf{i} + u\mathbf{j} + w\mathbf{k}$. We shall write this as $z + wj$ to emphasise that z and w are possibly not real. This gives us the useful laws

$$(z_1 + w_1 j)(z_2 + w_2 j) = (z_1 z_2 - w_1 \overline{w}_2) + (z_1 w_2 + w_1 \overline{z}_2)j, \quad jz = \overline{z}j, \quad (z + wj)\overline{(z + wj)} = |z|^2 + |w|^2.$$

Identify $\mathbb{C} \times \mathbb{R}$ with a 3-dimensional subspace of $\mathbb{H}$ via the embedding $(z, t) = (x + iy, t) \mapsto x\mathbf{1} + y\mathbf{i} + t\mathbf{j} = tj$; hence we can write

$$H^3 = \{z + tj \,:\, z \in \mathbb{C}, t \in \mathbb{R}_{>0}\} \subseteq \hat{\mathbb{R}}^3$$
$$\partial H^3 = \hat{\mathbb{C}}.$$

Now consider the action of $\mathbb{M}$ on H^3 via the Poincaré extension.

2.2.3 Lemma. *Suppose $f \in \mathbb{M}$ is of the form $f(z) = \frac{az+b}{cz+d}$, where we take a representative in* SL(2, $\mathbb{C}$)*; then $\tilde{f} : H^3 \to H^3$ has the form*

$$\tilde{f}(q) = (aq + b)(cq + d)^{-1}.$$

Proof. Recall (Proposition 1.2.14) that we may write f in the form

$$f(z) = \frac{a}{c} - \frac{ad - bc}{c^2}\left(z + \frac{d}{c}\right)^{-1}.$$

The Poincaré extension of f will be a similar formula, but with the reflections in $S(0, 1)$ and $y = 0$ replaced.

Indeed, if $q \in H^3$ is of the form $q = z + tj$ then the reflection in $S(0, 1)$ is $q^* := q/|q|^2 = (z + tj)/(|z|^2 + t^2)$, and the reflection in $y = 0$ of q^* is $(\overline{z} + tj)/(|z|^2 + t^2)$. Hence (assuming we have chosen the coordinates a, b, c, d so the corresponding matrix is in SL(2, $\mathbb{C}$))

$$\tilde{f}(q) = \frac{a}{c} - \frac{1}{c^2}\frac{\overline{z + d/c + tj}}{|z + d/c|^2 + t^2};$$

with routine algebra, one can rearrange this to obtain that

$$\tilde{f}(q) = \frac{(az + b)\overline{(cz + d)} + a\overline{c}t^2 + |ad - bc|tj}{|cz + d|^2 + |c|^2 t^2} = (a(z + tj) + b)\,(c(z + tj) + d)^{-1}. \qquad \square$$

We shall identify g and $\tilde{g}$ from now on.

2.2.4 Example. A warning: consider the transformation g given by $z \mapsto 1/z$. With the obvious matrix representation $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, note that using the lemma above we have that $g(j) = j^{-1} = -j$. Clearly this is nonsense, as we constructed the Poincaré extension such that preserved the sign of the jth component. The issue of course is that $\det A = -1$; to remedy this, we must multiply A through by the constant $(\det A)^{-1/2} = -i$ to obtain the representative $A' = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}$ which does have the correct determinant; then the lemma gives $g(j) = (-i)(-ij)^{-1} = (-i)(-k)^{-1} = -ik = j$, as expected (since the Poincaré extension of the circle inversion through the unit circle is the circle inversion through the unit sphere, and j lies on this sphere).

2.2.5 Example. Suppose $g \in \mathbb{M}$, represented by the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2, \mathbb{C})$, has the property $g(j) = j$. Then

$$\begin{aligned}(aj + b) &= j(cj + d) \\ &= \bar{c}j^2 + \bar{d}j \\ &= \bar{d}j - \bar{c}\end{aligned}$$

so $a = \bar{d}$ and $b = -\bar{c}$. In particular,

$$A^{-1} = \begin{bmatrix} \bar{d} & -\bar{c} \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} d & c \\ -\bar{c} & \bar{d} \end{bmatrix} = A^*$$

and $A \in \mathrm{SU}(2, \mathbb{C})$.

Conversely, if $A \in \mathrm{SU}(2, \mathbb{C})$ then A has the given form and thus, running the argument backwards, g fixes j.

2.2.6 Proposition. *If $g \in \mathbb{M}$, then $\|g\|^2 = 2\cosh \rho(j, g(j))$.*

Proof. Combine the formula

$$g(j) = \frac{b\bar{d} + a\bar{c} + j}{|d|^2 + |c|^2}$$

from the previous lemma with the expression from Theorem 1.5.7, to obtain

$$\begin{aligned}\cos \rho(j, g(j)) &= 1 + \frac{1}{2}(|d|^2 + |c|^2)\left\| j - \frac{b\bar{d} + a\bar{c} + j}{|d|^2 + |c|^2}\right\|^2 \\ &= 1 + \frac{1}{2(|d|^2 + |c|^2)}\left\|(|d|^2 + |c|^2 - 1)j - b\bar{d} - a\bar{c}\right\|^2 \\ &= 1 + \frac{1}{2(|d|^2 + |c|^2)}\left(\left|b\bar{d} + a\bar{c}\right|^2 + (|d|^2 + |c|^2)^2 - 2(|d|^2 + |c|^2) + 1\right);\end{aligned}$$

noting that $\left|b\bar{d} + a\bar{c}\right|^2 + 1 = (|a|^2 + |b|^2)(|c|^2 + |d|^2)$, we have

$$\begin{aligned}\cos \rho(j, g(j)) &= 1 + \frac{1}{2(|d|^2 + |c|^2)}\left((|a|^2 + |b|^2)(|c|^2 + |d|^2) + (|d|^2 + |c|^2)^2 - 2(|d|^2 + |c|^2)\right) \\ &= 1 + \frac{1}{2}\left(|a|^2 + |b|^2 + |d|^2 + |c|^2 - 2\right) \\ &= \frac{1}{2}\|g\|^2\end{aligned}$$

as desired. ☒

2.2.7 Theorem. *Suppose $g \in \mathbb{M}$ is represented by the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2, \mathbb{C})$. Let $f : H^3 \to B^3$ be the map of Eq.* (1.6.1)*. Then the following are equivalent:*

1. *$A \in \mathrm{SU}(2, \mathbb{C})$,*

2. *$g(j) = j$,*

3. *$||g||^2 = 2$,*

4. *$fgf^{-1} \in O(3)$, and*

5. *g is an isometry of $(\hat{C}, d)$ (where d is the chordal metric).*

Proof. The equivalence of (1) and (2) is Example 2.2.5. The equivalence of (2) and (3) is Proposition 2.2.6.

Now (2) is equivalent to $fgf^{-1}(0) = 0$, since f maps $e_{n+1} \mapsto 0$. Hence by Lemma 1.4.1 we are done.

Finally, recall the chordal metric is defined by

$$d(x,y) = \begin{cases} \dfrac{2||x-y||}{(1+||x||)^2)^{1/2}(1+||y||)^2)^{1/2}} & x,y \in \mathbb{C} \\ \dfrac{2}{1+||x||^2} & y = \infty. \end{cases}$$

Thus to show that g is an isometry is equivalent to showing that for all $z \in \mathbb{C}$,

$$\frac{|g'(z)|}{1+|g(z)|^2} = \frac{1}{1+|z|^2}.$$

Now we compute

$$\begin{aligned} \frac{|g'(z)|}{1+|g(z)|^2} &= \frac{|a(cz+d)-c(az+b)|}{|cz+d|^2\left(1+\left|\frac{az+b}{cz+d}\right|^2\right)} \\ &= \frac{|ad-bc|}{|cz+d|^2\left(1+\left|\frac{az+b}{cz+d}\right|^2\right)} \\ &= \frac{1}{|az+b|^2+|cz+d|^2} \end{aligned}$$

and so g is an isometry iff $|az+b|^2+|cz+d|^2 = 1+|z|^2$ for all $z \in \mathbb{C}$. Expanding we find $|az+b|^2 + |cz+d|^2 = (|a|^2+|c|^2)|z|^2 + (|b|^2+|d|^2) + 2\,\mathrm{Re}(a\overline{b}+c\overline{d})z$ and comparing coefficients we have $a\overline{b}+c\overline{d} = 0$ and $|a|^2+|c|^2 = |b|^2+|d|^2 = 1$; these equalities are equivalent to $A^* = A^{-1}$. This shows the equivalence of (1) and (5). ⅅ

2.2.8 Corollary. *The classical symmetry groups of the regular solids in B^3 are precisely the finite subgroups of* $\mathrm{SU}(2,\mathbb{C})$. ⅅ

2.3 Fixed points and conjugacy classes

2.3.1 Lemma. *A non-identity element of $\mathbb{M}$ has either one or two fixed points in $\hat{\mathbb{C}}$.*

Proof. Fix a representative $A \in \mathrm{SL}(2,\mathbb{C})$ for the transformation. Since we are working over $\mathbb{C}$, A has either one or two distinct eigenvalues. Thus we have three possibilities:

- A single eigenvalue with algebraic multiplicity 2 and geometric multiplicity 1: in this case, we have a single fixed line in $\mathbb{C}^2$ and thus a single fixed point in $\hat{\mathbb{C}}$.
- A single eigenvalue with algebraic multiplicity 2 and geometric multiplicity 2: in this case, we have that every element of $\mathbb{C}^2$ lies in this eigenspace, and A acts on $\hat{\mathbb{C}}$ as the identity.
- Two distinct eigenvalues each with algebraic multiplicity 1 and geometric multiplicity 1: in this case, we have a two distinct fixed lines in $\mathbb{C}^2$ and thus two fixed points in $\hat{\mathbb{C}}$.

As a consequence of this we get a simple proof that $\mathbb{M}$ is sharply triply transitive: if f and g both send $(w_1, w_2, w_3) \mapsto (z_1, z_2, z_3)$ then the composition fg^{-1} has three fixed points and thus is the identity.

We proceed to classify the elements of $\mathbb{M}$ according to their fixed points.

2.3.1 One fixed point

If $f \in \mathbb{M}$ has a unique fixed point, it is called **parabolic**.

2.3.2 Lemma. *Every parabolic element $f \in \mathbb{M}$ is conjugate to the translation $z \mapsto z + 1$.*

Proof. Let z_1 be the fixed point of f, let z_2 be any other point, and let $z_3 = f(z_2)$. Let g be the fractional linear map sending the triple $(z_1, z_2, z_3) \mapsto (\infty, 0, 1)$. Then fgf^{-1} has a unique fixed point at ∞ and maps 1 to 0. Recall that a fractional linear transformation fixing ∞ is of the form $z \mapsto az+b$, and such a map sends 1 to 0 iff $b = 1$. Further, $z = az+1 \implies (1-a)z = 1$ so such a transformation has a finite fixed point iff $a \neq 1$. Thus fgf^{-1} is the prototype translation.

Since trace is conjugation invariant, we see that $f \in \mathbb{M}$ is parabolic iff $\operatorname{tr}^2 f = 4$.

2.3.3 Lemma. *If g is parabolic with fixed point $x \in \hat{\mathbb{C}}$, then there is a unique non-zero $p \in \mathbb{C}$ such that the matrix*

- $\begin{bmatrix} 1 & p \\ 0 & 1 \end{bmatrix}$ *if $x = \infty$;*
- $\begin{bmatrix} 1+px & -px^2 \\ p & 1-px \end{bmatrix}$ *if $x \neq \infty$*

is a representative of g in $\mathrm{SL}(2, \mathbb{C})$.

The representative of Lemma 2.3.3 is the **normal form** of the element g.

Proof. If $x = \infty$ then g is of the form $g(z) = az + b$; since g has no finite fixed point we must have $a = 1$; and we may take $A = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$. On the other hand, suppose $x \neq \infty$. Since $\mathbb{M} \simeq \mathrm{PSL}(2, \mathbb{C})$ we may choose a unique representative matrix $A \in \mathrm{SL}(2, \mathbb{C})$ with trace 2 (the two representatives have traces ± 2 respectively), say $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $x = 0$ then $f(0) = b/d$ so $b = 0$ and thus $\det A = ad$, given that $ad = 1$ and $a + d = 2$ we have $a = d = 1$ and $f(z) = \frac{z}{cz+1}$ so we may take $p = c$. Conversely, suppose $x \neq 0$; then we may pick a unique p such that $a = 1 + px$ and $b = 1 - px$ (namely, take $p = (a-1)/x$ and then $b = 2 - a = 2 - (1 + px) = 1 - px$). Thus

$$1 = \det A = \det \begin{bmatrix} 1+px & b \\ c & 1-px \end{bmatrix} = 1 - p^2x^2 - bc$$

and so $bc = -p^2x^2$. Further, x is a fixed point of f and so

$$((1 + px)x + b) = (cx + (1 - px))x \implies (2p - c)x^2 = -b;$$

substituting we have $(2p - c)cx^2 = p^2x^2$, thus $0 = p^2 - 2pc + c^2 = (p - c)^2$ and $c = p$; this shows $b = -px^2$ as required.

2.3.2 Two fixed points

Consider the family $\mathscr{E}$ of maps $e_{k^2} \in \mathbb{M}$, indexed by $k \in \mathbb{C}^*$, $k \neq \pm 1$, defined by

$$e_{k^2}(z) := k^2 z.$$

Clearly each of these maps has exactly two fixed points, 0 and ∞. We call the value k^2 the **multiplier** of e_{k^2}, and note that $\operatorname{tr}^2 e_{k^2} = (k + k^{-1})^2$.

2.3.4 Lemma. *The map* tr^2 *sets up a 2-1 map from the space* $\mathscr{E}$ *to* $\mathbb{C} \setminus \{0, 4\}$, *with preimages being of the form* $\{k^2, k^{-2}\}$.

Proof. Suppose $t \in \mathbb{C} \setminus \{0, 4\}$. We wish to solve $t^2 = (k + k^{-1})^2$ for k^2. Rearranging, we obtain $0 = k^4 + (2 - t^2)k^2 + 1$ which is a quadratic polynomial in k^2 with discriminant $t^4 - 4t^2$. This polynomial in t^2 has solutions precisely when $t^2 \in \{0, 4\}$ which are precisely the disallowed values; thus each tr^2 comes from precisely two distinct values for k^2: if one is k^2, the other is k^{-2}.

By Lemma 2.3.4, conjugacy classes in $\mathscr{E}$ consist of at most two elements. Note that a given e_{k^2} has the property that, for all $z \in \mathbb{C}^*$,

$$\lim_{t\to\infty} \left| e^t_{k^2}(z) \right| = \left| (k^2)^t z \right| = \begin{cases} 0 & \left| k^2 \right| < 1 \\ \infty & \left| k^2 \right| > 1 \end{cases}$$

and so the fixed points of each map consist of one attractive and one repelling point; given any element e_{k^2}, the element $e_{k^{-2}}$ has the same trace but swapped nature of the fixed points. This suggests the following, which is too trivial even to be a lemma:

2.3.5 Observation. *The maps* e_{k^2} *and* $e_{k^{-2}}$ *are congruent in* $\mathbb{M}$.

Proof. Take the obvious map swapping 0 and ∞, namely $z \mapsto 1/z$. Conjugating by this map works.

If $\left|k^2\right| = 1$, we call e_{k^2} a **rotation**; if $k^2 \in (0, \infty) \setminus \{1\}$ then we call e_{k^2} a **dilation**.

2.3.6 Lemma. *The element* e_{k^2} *is a rotation if and only if* $\operatorname{tr}^2 e_{k^2} \in (0, 4)$. *The element* e_{k^2} *is a dilation if and only if* $\operatorname{tr}^2 e_{k^2} \in (4, \infty)$.

Proof. Note that $\left|k^2\right| = 1$ implies that $\operatorname{tr}^2 e_{k^2} = (k + \overline{k})^2 = 4(\operatorname{Re} k)^2$ and so $\operatorname{tr}^2 e_{k^2} \in [0, 4]$; and we are neglecting ± 1 so the endpoints of the interval are not attained. Similarly, $k^2 \in (0, \infty) \setminus \{1\}$ implies $\operatorname{tr}^2 e_{k^2} = (k + k^{-1})^2$ is a positive real number; if $k \neq 0$ then $\left|k + k^{-1}\right| \geq 2$ with equality exactly at $k = \pm 1$ (e.g. consider the local extrema of the hyperbola with equation $y = x + 1/x$).

For the converses, suppose $(k + k^{-1})^2 \in (0, \infty)$; in particular, $k + k^{-1}$ is real. Write $k = a + bi$; then we have $0 = \operatorname{Im}(k + k^{-1}) = b - b/(|a|^2 + |b|^2)$ and either $b = 0$ or $|k|^2 = |a|^2 + |b|^2 = 1$. Thus if $(k + k^{-1})^2 \in (0, \infty)$ then e_{k^2} is either a rotation or a dilation. We then use the first direction to see that if $\operatorname{tr}^2 e_{k^2} \in (0, 4)$ then e_{k^2} is a rotation, otherwise we obtain a contradiction; and similarly for the dilation case.

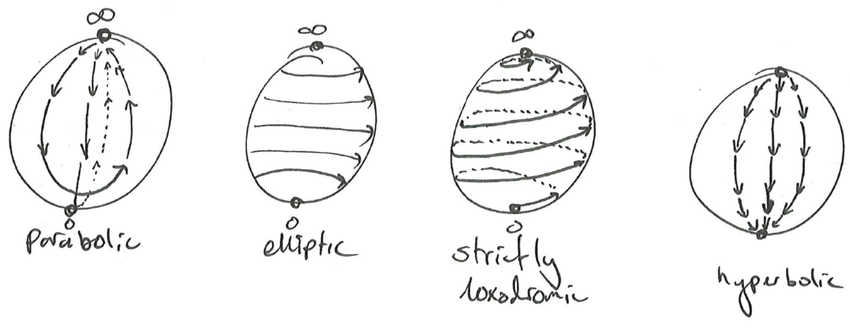


Figure 2.1: The shapes of the orbits of the different types of fractional linear transformation.

In particular, we see that the classes of rotations and dilations are disjoint. (Of course we could also see this by seeing that the map sending each e_{k^2} to its unique conjugate, namely $k^2 \mapsto k^{-2}$, preserves the classes.)

2.3.7 Definition. Let $g \in \mathbb{M}$ have precisely two fixed points in $\hat{\mathbb{C}}$. We call g variously:

- **elliptic**, if g is conjugate to a rotation;
- **hyperbolic**, if g is conjugate to a dilation;
- **strictly loxodromic**, if g is neither elliptic nor hyperbolic.

The class of **loxodromic** elements is the union of the classes of hyperbolic and strictly loxodromic elements.

We shall prove in a moment that every element with two distinct fixed points lies in one of these classes. The four possibilities for the type of an element of $\mathbb{M}$ are therefore parabolic, elliptic, strictly loxodromic, or hyperbolic. The names come from the shapes of the orbits; see Fig. 2.1.

2.3.8 Proposition. *An $g \in \mathbb{M}$ with two distinct fixed points in $\hat{\mathbb{C}}$ is conjugate to precisely two transformations of the form e_{k^2}.*

Proof. Let z_1 and z_2 be the two fixed points of g; pick any transformation $f \in \mathbb{M}$ sending $z_1 \mapsto 0$ and $z_2 \mapsto \infty$; the resulting map fgf^{-1} fixes 0 and ∞, hence is an element of $\mathscr{E}$. Then Observation 2.3.5 shows that fgf^{-1} has at least two elements of $\mathscr{E}$ in its conjugacy class, and Lemma 2.3.4 shows it has at most two.

In fact, the two conjugacy class representatives in $\mathscr{E}$ correspond to a choice of order of the fixed points of g. In the case that g is loxodromic it is possible to distinguish the fixed points via their flow properties; we shall always choose the representative with $|k^2| > 1$, so that the attractive fixed point is conjugated to ∞. If g is elliptic and $k^2 \neq -1$, we order the fixed points by choosing the representative k^2 in the upper half-plane (that is, the representative such that $k^2 = \exp(i\theta)$ with $0 < \theta < \pi$). If $k^2 = 1$ then the transformation $z \mapsto 1/z$ swaps $0 \leftrightarrow \infty$ and commutes with -1, so we cannot intrinsicly distinguish between the fixed points; if one of the fixed points is ∞ we distinguish that one, otherwise we will always choose $k = i$ as the distinguished conjugate.

2.3.9 Corollary. *Two elements $f, g \in \mathbb{M}$ are conjugate if and only if* $\operatorname{tr}^2 f = \operatorname{tr}^2 g$.

Proof. One direction was Lemma 2.2.1. It remains to show that $\operatorname{tr}^2 f = \operatorname{tr}^2 g$ implies $f \sim g$. If $\operatorname{tr}^2 f = 4$, this was the content of the previous section (parabolic elements form a conjugacy class of $\mathbb{M}$ and are thus characterised by their tr^2, which is 4). Otherwise, f and g have two fixed points. If $\operatorname{tr}^2 f = \operatorname{tr}^2 g$, then both f and g are conjugate to the same element of $\mathscr{E}$ (since conjugacy classes of $\mathscr{E}$ are precisely given by the value of tr^2, by Observation 2.3.5 and Lemma 2.3.4) and we are done. ◻

We now turn to the problem of finding a normal form for the transformations with two fixed points.

2.3.10 Lemma. *Let g be a transformation with fixed points x and y; order the fixed points so that the distinguished fixed point is y. Then the matrix*

$$A = \begin{cases} \frac{1}{x-y}\begin{bmatrix} xk^{-1} - yk & xy(k-k^{-1}) \\ k^{-1} - k & xk - yk^{-1} \end{bmatrix} & x, y \neq \infty \\ \begin{bmatrix} k^{-1} & y(k-k^{-1}) \\ 0 & k \end{bmatrix} & x = \infty \\ \begin{bmatrix} k & x(k^{-1} - k) \\ 0 & k^{-1} \end{bmatrix} & y = \infty \end{cases}$$

is a representative for g in $\mathrm{SL}(2,\mathbb{C})$*; this is the* ***normal form*** *of g.*

Proof. It suffices to check that y is the distinguished fixed point of A in each case, and that $g(0) = A0$, $g(1) = A(1)$, and $g(\infty) = A(\infty)$. These are routine calculations. ◻

The difference between hyperbolic and strongly loxodromic elements may also be described in terms of invariant discs. We note that when we say **disc** we will always mean either the interior of a Euclidean circle, or a half-plane.

2.3.11 Proposition. *Let g be a loxodromic element. Then g is hyperbolic iff there is an open disc D in* $\hat{\mathbb{C}}$ *left invariant by g; otherwise g is strongly loxodromic.*

Proof. Suppose g is hyperbolic; conjugate g via f to some element $z \mapsto \lambda^2 z$ with $\lambda^2 > 1$. This element fixes the upper half-plane of $\hat{\mathbb{C}}$, and so g fixes the inverse image of this half-plane under f (which is the interior of a circle). Conversely, if g is conjugate to some element $z \mapsto r\exp(i\theta)z$ with $\theta \neq 0$ then this conjugate fixes no Euclidean disc (discs centred at 0 are mapped to strictly larger discs, and discs not centred at 0 are rotated off themselves) and no half-space (either the boundary line is rotated off itself, or the half-spaces bounded by it are exchanged if the line passes through 0 and $\theta = \pi$). ◻

2.3.3 The trace of an element

We now have the necessary results to completely classify the conjugacy classes of $\mathbb{M}$.

2.3.12 Proposition. *Let* $g \in \mathbb{M}$ *be an arbitrary element. Then:*

- $\operatorname{tr}^2(g)$ *is real and lies in* $[0,4)$ *if and only if g is elliptic;*
- $\operatorname{tr}^2(g) = 4$ *if and only if f is parabolic or the identity;*
- $\operatorname{tr}^2(g)$ *is real and lies in* $(4,\infty)$ *if and only if g is hyperbolic;*
- $\operatorname{tr}^2(g)$ *does not lie in the segment* $[0,4]$ *if and only if g is loxodromic.*

Proof. Compute with the normal forms and use that tr^2 is invariant under conjugacy. ◻

We also give an alternative classification.

2.3.13 Proposition. *Let $g \in \mathbb{M}$ be an arbitrary non-identity element. Then g has at least one fixed point in $\overline{H^3} = H^3 \cup \hat{\mathbb{C}}$; in the following we consider the action of g on this set.*

- *g is elliptic $\iff$ g has a fixed point in H^3 $\iff$ g has infinitely many fixed points $\iff$ the fixed points of g are precisely the points of the closure of a hyperbolic line (called the* ***axis*** *of g).*
- *g is parabolic $\iff$ g is not elliptic, and has exactly one fixed point in $\hat{\mathbb{C}}$.*
- *g is loxodromic $\iff$ g is neither elliptic nor parabolic $\iff$ g has exactly two fixed points.*

Remark. We take these to be the *definition* of the adjectives parabolic/loxodromic/elliptic when we consider groups of isometries of $\mathbb{H}^n$ for $n > 2$. Note that while some of these conditions are almost trivial when we consider $g \in \mathbb{M}$, in the general case they are not (e.g. in general, elements might have more than two fixed points on ∂H^3). Compare [34, section IV.C] and [6, definition 4.3.2].

Proof of Proposition 2.3.13.

- Suppose g is elliptic; let it act on the half-space model of H^3, and conjugate to the element $z \mapsto k^2 z$ for $\left|k^2\right| = 1$. Then the Poincaré extension of g clearly has as fixed points exactly those points on the line $x_1 = x_2 = 0$. This shows both that g has a fixed point in H^3, and that it has infinitely many fixed points altogether, on a hyperbolic line.

  Now suppose g has infinitely many fixed points; since it has only two fixed points on $\hat{\mathbb{C}}$, it must have at least one fixed point in H^3 (here we use the deep result that $\infty > 2$).

  Finally, we show that having a fixed point in H^3 implies both having infinitely many fixed points and being elliptic. Consider the action of g on the ball model of H^3, conjugated so that the fixed point in H^3 is 0. By Lemma 1.4.1, $g \in O(3)$. The fixed points of g are the points $x \in B^3$ such that $gx = 1x$, i.e. the nullspace of $(1 - g)$. This is a Euclidean subspace, in fact it must be a line through the origin (it has either 1 or 2 intersection points with ∂B^3 and every flat has at least 2 such intersection points if it has any); in particular, g fixes pointwise a hyperbolic line. The orbits of ∂B^3 are latitudes with respect to this axis, and in particular g is an element with two fixed points which are not attractive or repelling; thus g is elliptic.

- Suppose g is parabolic; then g has not elliptic and has exactly one fixed point in $\hat{\mathbb{C}}$ by definition. Conversely, suppose g is not elliptic; then g is either loxodromic or parabolic by definition, and if it has exactly one fixed point in $\hat{\mathbb{C}}$ then it must be parabolic.

- The equivalence g loxodromic $\iff$ g is neither elliptic nor parabolic is by consideration of the trace, as above. Suppose g has exactly two fixed points; note that g cannot have any fixed points in H^3, for then it would fix an entire hyperbolic line (as in the elliptic case) and so g has exactly two fixed points in $\hat{\mathbb{C}}$. Since it is not elliptic, it must be loxodromic. Conversely, if g is loxodromic then, in particular, it is not elliptic, so it does not fix any points of H^3, and thus has exactly two fixed points.

A stronger version of Proposition 2.3.11. Recall that we use *disc* to mean a Euclidean disc (i.e. interior or exterior of a circle) or a half-plane in $\hat{\mathbb{C}}$.

2.3.14 Proposition. *An element $g \in \mathbb{M}$ leaves a disc invariant iff $\operatorname{tr}^2 g \geq 0$.*

Proof. Suppose $\operatorname{tr}^2 g \geq 0$. By Proposition 2.3.12, we have that g is one of the identity, elliptic, parabolic, or hyperbolic. Clearly the identity fixes a disc. Parabolic transformations are conjugate to $z \mapsto z + 1$, which fixes the upper half-plane (in fact, any half-plane bounded by a horizontal line).

Elliptic transformations are conjugate to $z \mapsto k^2 z$ for $|k^2| = 1$, every transformation of this type fixes the unit disc (in fact, any disc about 0). Hyperbolic transformations are conjugate to $z \mapsto k^2 z$ for $k^2 \in \mathbb{R}_{>0}$, which fixes the upper half-plane (in fact, any half-plane bounded by a line through 0).

Conversely, suppose $\operatorname{tr}^2 g \notin [0, \infty)$. Then g is strictly loxodromic, and we proved this case as Proposition 2.3.11.

In fact, note that if $g \in \mathbb{M}$ has $\operatorname{tr}^2 g \geq 0$ and if $z_0 \in \hat{\mathbb{C}}$ is not a fixed point of g, then conjugating g to be of a regular form (either $z \mapsto z + 1$ or $z \mapsto k^2 z$) sends z_0 to a point distinct from ∞ and 0; in particular, there exists a disc bounded by a circle S through z_0 such that the interior and exterior discs of S are left invariant by g.

2.4 Commutators and fixed points

In this section, we shall study the behaviour of pairs of elements $g, h \in \mathbb{M}$ which share a fixed point in $\hat{\mathbb{C}}$ (Lemma 2.4.2) or in H^3 (Theorem 2.4.7). See [6, pp. 68–74], as well as [34, p. I.D].

Let G be a group acting on some set X. For $g \in G$, write $\operatorname{Fix}_X g$ for the set of all $x \in X$ such that $gx = x$. Recall also that the **commutator** of $g, h \in G$ is the element $[g, h] \coloneqq ghgh^{-1}h^{-1}$ (so $[g, h] = 1 \iff gh = hg$). Note that if $g, h \in \mathbb{M}$ have representatives $A, B \in \mathrm{SL}(2, \mathbb{C})$ then $[g, h] = [A, B]$ (so the commutator is independent of the choice of matrix representatives).

2.4.1 Lemma. *If $g, h \in \mathbb{M}$ act on $\hat{\mathbb{C}}$, then g and h have a common fixed point iff* $\operatorname{tr}[g, h] = 2$.

Proof. Suppose g, h have a common fixed point, x; let $f \in \mathbb{M}$ send $x \mapsto \infty$. Conjugation leaves trace invariant, so it suffices to check that $\operatorname{tr}[g^f, h^f] = 2$; if some element $A \in \mathrm{PSL}(2, \mathbb{C})$ fixes ∞, then the lower-left entry of A is zero. A simple computation shows that if A and B are two matrices with this property, then $ABA^{-1}B^{-1}$ has trace 2.

Conversely, suppose g and h have $\operatorname{tr}[g, h] = 2$; we may assume that g fixes ∞, and so the matrix representatives for g and h are respectively $A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ and $B = \begin{bmatrix} x & y \\ u & v \end{bmatrix}$. Using that $A, B \in \mathrm{SL}(2, \mathbb{C})$ we have

$$2 = \operatorname{tr}[A, B] = b^2u^2 - bu(a - d)(v - x) + 2vx - (a^2 + d^2)uy;$$

in the case that g is parabolic, by Lemma 2.3.3 we have $a = d = 1$ and $b \neq 0$, so

$$2 = b^2u^2 + 2vx - 2uy = b^2u^2 + 2$$

so $b^2u^2 = 0$ and $u = 0$, so h fixes ∞. On the other hand, if g is not parabolic we may conjugate again to ensure that the second fixed point of g is 0, and so by Lemma 2.3.10 we have $b = 0$ and $a \neq d$. In particular, using $vx - uy = 1$,

$$2 = 2vx - (a^2 + d^2)uy = 2 + 2uy - (a^2 + d^2)uy \implies 0 = (a^2 - 2ad + d^2) = uy = (a - d)^2 uy;$$

since $a \neq d$, $uy = 0$ and so either $u = 0$ and $h(\infty) = \infty$, or $y = 0$ and $h(0) = 0$.

2.4.2 Lemma. *If $g, h \in \mathbb{M}$ are nontrivial and have a common fixed point in $\hat{\mathbb{C}}$, then either*

1. $[g, h] = 1$ *and* $\operatorname{Fix} g = \operatorname{Fix} h$*; or*

2. $[g, h]$ *is parabolic and* $\operatorname{Fix} g \neq \operatorname{Fix} h$.

Proof. Again, if g and h have a common fixed point we may conjugate (which preserves the commutator and the type of the transformation) so that ∞ is a common fixed point. Note that $\operatorname{tr}[g,h] = 2$ by Lemma 2.4.1, so either $[g,h] = 1$ or $[g,h]$ is parabolic; a computation shows that

$$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}\begin{bmatrix} x & y \\ 0 & v \end{bmatrix}\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}^{-1}\begin{bmatrix} x & y \\ 0 & v \end{bmatrix}^{-1} = \begin{bmatrix} 1 & \frac{bv+ay-bx-dy}{dv} \\ 0 & 1 \end{bmatrix},$$

so $[g,h] = 1$ iff $bv + ay - bx - dy = 0$ iff $b(v-x) = y(a-d)$: if $a = d$, then $b = 0$ and by the normal forms if $b = 0$ then either g is parabolic (if $a \neq d$) or the identity (if $a = d$); neither of these cases are possible here, so we must have $a \neq d$ so g has two fixed points. Conjugate again so that g has fixed points at 0 and ∞, then we have

$$\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}\begin{bmatrix} x & y \\ 0 & v \end{bmatrix}\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}^{-1}\begin{bmatrix} x & y \\ 0 & v \end{bmatrix}^{-1} = \begin{bmatrix} 1 & \frac{y(a-d)}{dv} \\ 0 & 1 \end{bmatrix},$$

and thus, since this is the identity matrix and $a \neq d$, $y = 0$ so h has fixed points at 0 and ∞.

Conversely, suppose the fixed points of g and h differ, conjugating the fixed points of g to 0 and ∞ and computing the commutator as above shows that $[g,h] \neq 1$.

2.4.3 Theorem. *Let $g, h \in \mathbb{M}$ be nontrivial. The following are equivalent:*

1. $[g,h] = 1$ *(g and h commute);*
2. $g(\operatorname{Fix}_{\hat{\mathbb{C}}} h) = \operatorname{Fix}_{\hat{\mathbb{C}}} h$ *and* $h(\operatorname{Fix}_{\hat{\mathbb{C}}} g) = \operatorname{Fix}_{\hat{\mathbb{C}}} g$*;*
3. *Either:*
   - *(a)* $\operatorname{Fix}_{\hat{\mathbb{C}}} h = \operatorname{Fix}_{\hat{\mathbb{C}}} g$*, or*
   - *(b) g and h have a common fixed point in H^3, no common fixed point in $\hat{\mathbb{C}}$, and the relations $g^2 = h^2 = (gh)^2 = 1$ hold;*
4. *Either:*
   - *(a)* $\operatorname{Fix}_{\hat{\mathbb{C}}} h = \operatorname{Fix}_{\hat{\mathbb{C}}} g$*, or*
   - *(b) g and h are elliptic of order 2, and each exchanges the fixed points of the other.*

Proof.

(1) $\implies$ (2). Suppose $gh = hg$. Let $x \in \operatorname{Fix}_{\hat{\mathbb{C}}} h$, then $gx = ghx = hgx$, so $gx \in \operatorname{Fix}_{\hat{\mathbb{C}}} h$; hence $g(\operatorname{Fix}_{\hat{\mathbb{C}}} h) \subseteq \operatorname{Fix}_{\hat{\mathbb{C}}} h$. Converse inclusion is easy (consider g^{-1}).

(2) $\implies$ (3). Suppose the closure properties of (2) hold, and that $\operatorname{Fix}_{\hat{\mathbb{C}}} h \neq \operatorname{Fix}_{\hat{\mathbb{C}}} g$. Pick some $x \in \operatorname{Fix}_{\hat{\mathbb{C}}} g \setminus \operatorname{Fix}_{\hat{\mathbb{C}}} h$. Then $x, hx, h^2x \in \operatorname{Fix}_{\hat{\mathbb{C}}} g$ since the latter is closed under left-multiplication by h. These cannot be distinct (h has at most two fixed points), and $hx \neq x$ (by assumption) so $x = h^2x$. In particular, g has exactly two fixed points (h and hx) which are exchanged by h. Conjugating, we may assume the fixed points of g are 0 and ∞, so $gz = az$ ($a \in \mathbb{C}$); since h exchanges 0 and ∞, we have $hz = b/z$ ($b \in \mathbb{C}$). Then, for $z \in \mathbb{C}$, $h^2z = z$ and $(gh)^2z = z$. This gives the relations. Finally, note that g and h both fix the point $|b|^{1/2}j$ (by direct substitution into Lemma 2.2.3: note that the correct form for g is $q \mapsto (\frac{a}{|a|^{1/2}}q)\frac{1}{|a|^{1/2}}$ etc.).

(3) $\implies$ (4). Suppose g and h have a common fixed point in H^3, in particular by Proposition 2.3.13 both are elliptic. Suppose x is a fixed point of g, then $ghx = h^2ghx = hgx = hx$ (since $ghgh = 1 \iff hgh = g$) and hx is the other fixed point; this shows the exchange property.

(4) $\implies$ (1). If $\operatorname{Fix}_{\hat{\mathbb{C}}} h = \operatorname{Fix}_{\hat{\mathbb{C}}} g$ then Lemma 2.4.2 allows us to conclude that $gh = hg$. On the other hand, suppose g and h are elliptic of order 2, and each exchanges the fixed points of the other. By conjugation, we may assume $g(z) = k^2 z$ for some $k \in \mathbb{C}$; this uses 2-transitivity. Using the further rigidity condition of 3-transitivity, conjugate one of the fixed points of h to 1; so h interchanges 0 and ∞ and fixes 1, so $h(z) = 1/z$. The other fixed point of h is -1, so g exchanges ± 1, i.e. $k^2 = -1$. Hence $f(z) = 1/z$ and $g(z) = -z$, and these transformations commute.

These two theorems essentially complete the study of common fixed points in $\hat{\mathbb{C}}$. We now move to the technically harder study of common fixed points in H^3.

Recall the notion of the **axis** of an elliptic element from Proposition 2.3.13. We write A_g for the axis of an elliptic element g. Note that $g, h \in G$ have a common fixed point in H^3 iff $A_g \cap A_h \cap H^3 \neq \emptyset$; this is equivalent to requiring the fixed points of g and h to lie on some circle Q, and to separate each other on Q (picture).

2.4.4 Lemma. *If g, h, gh are all elliptic, then the fixed points of g and h in $\hat{\mathbb{C}}$ are concyclic. Further, if $[g, h]$ is either elliptic or 1, then A_g and A_h intersect in H^3.*

Proof. We split into two cases: that g and h have a common fixed point in $\hat{\mathbb{C}}$, and that they do not.

g and h have a common fixed point. In this case, $\operatorname{Fix}_{\hat{\mathbb{C}}} g \cup \operatorname{Fix}_{\hat{\mathbb{C}}} h$ has at most three points, so determines a circle — this proves concyclicity. Now suppose $[g, h]$ is either elliptic or 1. Considering the cases in the conclusion of Lemma 2.4.2, we must actually have that $[g, h] = 1$ (as two noncommuting elements with common fixed points must have a parabolic commutator and this is not possible here by assumption) and so the fixed points of g and h coincide; hence the axes coincide (and have infinitely many points of intersection).

g and h have no common fixed point. For the second case, assume g and h have no common fixed points, and conjugate the fixed points of g to 0 and ∞; thus we have $g(z) = k^2 z$ for some $k^2 \in \mathbb{C} \setminus \{1\}$ with $|k^2| = 1$. Pick a matrix representative for h, say $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2, \mathbb{C})$. By direct computation, $\operatorname{tr}^2 h = (a + d)^2$ and $\operatorname{tr}^2(gh) = (ka + \bar{k}d)^2$; each of these elements lies in $[0, 4)$, so $\lambda = a + d$ and $\mu = ka + \bar{k}d$ lie in $(-2, 2)$. We have $d = \frac{\mu - \lambda k}{\bar{k} - k}$ and $a = \bar{d}$. Write $a = u + iv$. Using this, we compute that the fixed points of h are

$$\zeta, \xi = \frac{i}{c}\left(v \pm (1 - u^2)^{1/2}\right); \tag{2.4.5}$$

since $|a + d| < 2$, we have $u^2 = \frac{1}{4}(a + d)^2 < 1$ and so ζ and ξ are both on the line L generated by i/c.

Suppose now that $[g, h]$ is elliptic or 1; we have $0 \leq \operatorname{tr}^2[g, h] < 4$, and a further computation gives $\operatorname{tr}^2[g, h] = 4\left(1 + (|a|^2 - 1)\sin^2\theta\right)^2$ where $\theta = \arg k$, so $|a|^2 - 1 < 0$. If $|a| = 1$ then $u^2 + v^2 = 1$ so $v \pm (1 - u^2)^{1/2}$ takes on the value zero; this contradicts the assumption that g, h have no common fixed points and so $|a| < 1$. i.e. $(1 - u^2)^{1/2} > v$. From Eq. (2.4.5) above, we may write $\zeta = is/c$ and $\xi = it/c$ for $s, t \in \mathbb{R}$ and one of s and t is negative; hence the two fixed points of h lie on opposite sides of 0 on the line L, so separate the fixed points of g, and therefore A_h and A_g intersect.

We now apply Lemma 2.4.4 to the study of subgroups of $\mathbb{M}$ generated by two elliptic elements.

2.4.6 Lemma. *Let $g, h \in \mathrm{GM}(3)$ be nontrivial elements of $\mathbb{M}$ acting B^3, fixing the origin (so every nontrivial element of $\langle g, h\rangle$ is elliptic). Then either*

- *the elements of $\langle g, h\rangle$ have the same axis and the same fixed points, or*
- *there is some $f \in \langle g, h\rangle$ such that A_g, A_h, A_f are not coplanar.*

Proof. By assumption, the axes A_g and A_h are diameters of B_3; assume $A_h \neq A_g$, so A_h and A_g determine a Euclidean plane P. Let D denote the (unique) diameter of B^3 orthogonal to P. If $h(A_g)$ does not lie in P, then $f = hgh^{-1}$ has the property that $A_f = h(A_g)$; similarly, if $g(A_h)$ does not lie in P, then take $f = ghg^{-1}$. Assume therefore that $g(A_h)$ and $h(A_g)$ both lie in P. This shows that g and h preserve P, and so act as reflections across it; they therefore exchange the endpoints of D and so $f = gh$ has axis D. ∎

2.4.7 Theorem. *Let $G \leq \mathbb{M}$. The elements of G have a common fixed point in H^3 iff all the non-identity elements of G are elliptic.*

Proof. It is trivial that if all the nontrivial elements of G have a shared fixed point in H^3 then they are elliptic. The converse is slightly harder. Suppose every nontrivial element of G is elliptic; we may assume that G contains two elements g, h such that $A_g \neq A_h$ (otherwise all the elements fix a common axis and have infinitely many shared fixed points). By Lemma 2.4.4, g and h have a common fixed point in H^3; let G act on B^3 and by conjugation send this common fixed point to 0. Note that the hypotheses of Lemma 2.4.6 hold, and by assumption we are the second case of the conclusions: there is some $f \in \langle g, h\rangle$ such that A_g, A_h, A_f are not coplanar; and since every element of $\langle g, h\rangle$ fixes 0, all of A_g, A_f, A_h are Euclidean diameters of B^3.

Pick any nontrivial $q \in G$. By Lemma 2.4.4, the fixed points of q and g in $\hat{\mathbb{C}}$ are concyclic; let Π_g be the Euclidean plane containing this circle. Since Π_g contains the endpoints of A_q and is a plane, $A_q \subseteq \Pi_g$; similarly, Π_g contains A_g and hence 0. Define similar planes Π_h and Π_f, so $0 \in \Pi_g \cap \Pi_h \cap \Pi_f$ and $A_q \subseteq \Pi_g \cap \Pi_h \cap \Pi_f$.

Note that Π_g, Π_h, Π_f are distinct, otherwise A_g, A_h, A_f would be colinear. Hence the intersection $\Pi_g \cap \Pi_h \cap \Pi_f$ is 0- or 1-dimensional, so is either 0 or A_q. In particular, $0 \in A_q$ and so $q(0) = 0$: 0 is a shared fixed point of G. ∎

2.4.8 Corollary. *If $G \leq \mathbb{M}$ is finite, then the elements of G have a common fixed point in H^3.*

Proof. If G is finite, each element $g \in G$ has finite order and so is elliptic; apply Theorem 2.4.7. ∎

Chapter 3

Kleinian groups

3.1 Discontinuous actions and the ordinary set

Fix a topological space X and let G be a group acting on X via homeomorphisms.

We say that the action is:

- **freely discontinuous** (or **properly discontinuous**, or a **covering space action**) on X if, for every $x \in X$, there exists a neighbourhood $U \ni x$ (called a **nice neighbourhood**) such that $gU \cap U = \emptyset$ for all $g \in G$ nontrivial.

- **proper** if the continuous map $\Theta : G \times X \to X \times X$ defined by $\Theta(g, x) = (x, gx)$ is proper.

- **openly discontinuous** if for every $x, y \in X$ there exist neighbourhoods U of x and V of y such that $U \cap gV = \emptyset$ for all but finitely many $g \in G$;

- **weakly discontinuous** if for every $x \in X$ there exists a **weakly nice neighbourhood** U of x such that $U \cap gU = \emptyset$ for all but finitely many $g \in G$;

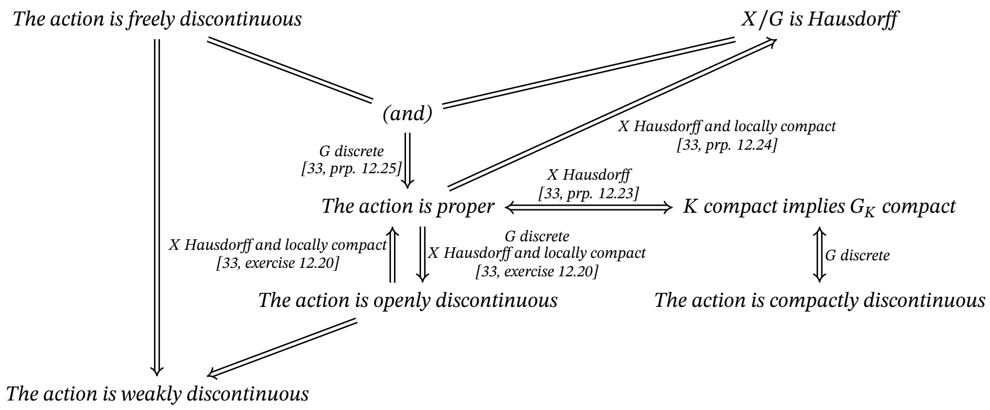
- **compactly discontinuous** if for every $K \subseteq X$ compact, $K \cap gK = \emptyset$ for all but finitely many $g \in G$. If $K \subseteq X$ is compact, define $G_K := \{g \in G : K \cap gK \neq \emptyset\}$; then compact discontinuity of the action is equivalent to $|G_K| < \infty$ for all K compact.

We write $^\circ\Omega(G)$ for the largest subspace of X on which the restriction of the group action is freely discontinuous (the **free regular set**), and we write $\Omega(G)$ for the largest subspace of X on which the restriction of the group action is weakly discontinuous (the **regular set**). It is easy to see that both of these subspaces are open in X. If the group G is understood from context, we shall merely write $^\circ\Omega$ and Ω.

The first set of relationships between the different concepts of discontinuity is as follows:

3.1.1 Theorem. *Let X be a topological space, and let G be a group acting on X by homeomorphisms.*

Then we have the following relationships:



Proof. References are given in the diagram, except for (1) the equivalence of 'K compact implies G_K' and 'compactly discontinuous' when G is discrete — this is immediate from the fact that compact subspaces of a discrete space are exactly the finite subspaces — and (2) the implication from 'openly discontinuous' to 'weakly discontinuous' which follows when $x = y$ and the neighbourhood of the definition is taken to be $U \cap V$.

We now reduce to the cases of interest: subgroups of $\mathbb{M}$ acting on $\hat{\mathbb{C}}$ and H^3. Both of these spaces are Hausdorff and locally compact. We will also often restrict ourselves to the case that G is discrete; with this setup, the notions of openly, and compactly discontinuous actions are equivalent and if the action has these properties then we say the action is **discontinuous**. It will follow from Proposition 3.1.4 below that, in the case of $\mathbb{M}$ acting on $\hat{\mathbb{C}}$, this is also equivalent to weak discontinuity.

We start with the action on H^3; the characterisation is easy:

3.1.2 Theorem. *A subgroup $G \leq \mathbb{M}$ is discrete iff it acts discontinuously in H^3.*

We shall need a lemma.

3.1.3 Lemma. *Discrete subgroups of* $\mathrm{SL}(2,\mathbb{C})$ *are countable. In fact, $G \leq \mathrm{SL}(2,\mathbb{C})$ is discrete iff for all $k \in \mathbb{N}$, the set $G_k = \{A \in G : \|A\| \leq k\}$ is finite.*

Proof. Clearly the countability statement follows from the second statement: $G = \bigcup_{k\in\mathbb{N}} G_k$ so is a countable union of finite sets.

If each G_k is finite, then G cannot have any limit points as $\|\cdot\|$ is continuous. Conversely, if some G_k is infinite then pick some sequence of distinct elements $(A_n) \subseteq G_k$; if we write $(A_n)_{ij}$ for the coordinates of A_n, note that $\left\|(A_n)_{ij}\right\| \leq k$ for all i, j and so (since $[-k, k]$ is compact) there exists a convergent subsequence of the (A_n), say converging to some $A \in \mathrm{End}(2,\mathbb{C})$. Since det is continuous, we actually have $A \in \mathrm{SL}(2,\mathbb{C})$; hence G has a limit point as a subspace of $\mathrm{SL}(2,\mathbb{C})$.

Proof of Theorem 3.1.2. Suppose G is discrete. By Lemma 3.1.3, G is countable (it is a homomorphic image of $\mathrm{SL}(2,\mathbb{C})$), say $\{g_1, ...\}$. By discreteness, $\|g_i\| \to \infty$. Now use Proposition 2.2.6 to see that $\rho(j, g_n(j)) \to \infty$ as $n \to \infty$. Now note that a compact subset $K \subseteq H^3$ lies in a hyperbolic ball B around j (indeed, one can show that hyperbolic balls in the half-plane model are Euclidean balls, though of different centre and radius), say of radius k. If $gK \cap K \neq \emptyset$, then $gB \cap B \neq \emptyset$ and so $\rho(j, g(j)) < 2k$. Since this quantity tends to ∞, there can be only finitely many g such that $gK \cap K \neq \emptyset$. This shows discontinuity.

If G is not discrete, pick distinct matrices $A_1, A_2, \ldots \in \mathrm{SL}(2,\mathbb{C})$ such that the images $g_1, \ldots \in G$ converge to 1. Note then that $g_n(x) \to x$ as $n \to \infty$ for all $x \in H^3$; in particular, the orbit Gx has a limit point in H^3, which implies G cannot act discontinuously (a limit point must be contained in infinitely many translates of a neighbourhood of x).

The theory is richer in the case of an action on $\hat{\mathbb{C}}$.

3.1.4 Proposition. *If the action of $G \leq \mathbb{M}$ on a G-invariant space $X \subseteq \hat{\mathbb{C}}$ is weakly discontinuous, then it is in fact openly discontinuous.*

Proof. Let $x, y \in X$; using weak discontinuity, find weakly nice neighbourhoods $U \ni x$ and $V \ni y$; without loss of generality, assume U and V are open discs centred at x and y respectively. There are at most finitely many translates of x in V (since V is a weakly nice neighbourhood of each gx lying in V); we may therefore shrink V such that no translate of x lies in $\overline{V}$.

Suppose there exists a sequence $(g_i) \subseteq G$ such that $g_i V \cap U \neq \emptyset$ for all i. Note that $g_i V \cap g_j V \neq \emptyset$ implies that $V \cap g_i^{-1} g_j V \neq \emptyset$; by weak discontinuity, there are only finitely many pairs i, j such that $g_i V$ intersects $g_j V$. Since all of these translates are subspaces of the 2-sphere $\hat{\mathbb{C}}$, we must have that the spherical diameter (i.e. the diameter obtained by pulling back the chordal metric as in Eq. (1.2.9)) $\operatorname{diam} g_i V \to 0$ as $i \to \infty$. In particular, the sequence of images of y must have an accumulation point on the boundary of U. But the translates of y by the g_i in X cannot have any accumulation points — if they accumulated near some point $z \in X$, then that point could not have a weakly nice neighbourhood. We may therefore conclude that there are only finitely many $g \in G$ such that $gV \cap U \neq \emptyset$.

Proposition 3.1.4, combined with Theorem 3.1.1, gives the following pair of corollaries:

3.1.5 Corollary. *Let $G \leq \mathbb{M}$ act on $\hat{\mathbb{C}}$. Then the orbit space ${}^\circ\Omega(G)/G$ is Hausdorff.*

3.1.6 Corollary. *If $G \leq \mathbb{M}$ is discrete and has freely discontinuous action on some $X \subseteq \hat{\mathbb{C}}$, then it is discontinuous on X.*

We call a discrete subgroup $G \leq \mathbb{M}$ a **Kleinian group**. If ${}^\circ\Omega(G) = \emptyset$, we say that G is **of the first kind**; otherwise, we say G is **of the second kind**. If $z \in \overline{H^3}$ is a fixed point of a parabolic element of G, we say that z is a **cusp** of G; if z is a fixed point of an elliptic, loxodromic, or hyperbolic element we say that z is variously an **elliptic point**, a **loxodromic point**, or a **hyperbolic point** of G.

Remark. Note that in [34], a *Kleinian group* is our Kleinian group of the second kind; in [6], a *Kleinian group* is a discrete subgroup $G \leq \mathbb{M}$ where $\Omega(G) \neq \emptyset$. We shall prove later (Corollary 3.3.26) that $\Omega(G) \neq \emptyset \iff {}^\circ\Omega(G) \neq \emptyset$ for such G, and so this also corresponds to a Kleinian group of the second kind. Significant historical notes may be found in [38, §5.6].

3.1.7 Proposition. *If $G \leq \mathbb{M}$ is an arbitrary subgroup with $\Omega(G) \neq \emptyset$, then G is discrete (hence Kleinian).*

Proof. Suppose G is not discrete; then there is a sequence (g_i) of elements of G such that $g_i \to 1$. Then, for all $z \in \hat{\mathbb{C}}$, $g_i(z) \to z$; so z is an accumulation point of the orbit Gz. Hence either $gz = z$ for infinitely many $g \in G$, or every neighbourhood of z contains infinitely many translates of z. In either case, every neighbourhood of z has the property that $U \cap gU \neq \emptyset$ for infinitely many g, and so $z \notin \Omega(G)$. In particular, $\Omega(G) = \emptyset$.

This proposition, as well as the proof of Proposition 3.1.4 above, suggests that it will be profitable to study the limit points of orbits of Kleinian groups G. We do this in the next section, but first we give some examples.

3.1.8 Example (Dihedral groups). The dihedral groups are precisely the noncyclic finite Kleinian groups with a cyclic normal subgroup. If $v \geq 2$ is an integer, let H be the cyclic subgroup generated by the rotation $z \mapsto \exp(2\pi i/v)z$, and let $G = \langle H, b\rangle$ where b is complex inversion. Note that b normalises H: $(\exp(2\pi i m/v)z)^{-1} = \exp(-2\pi i m/v)z^{-1}$, so $bH = Hb$. Conclude that H is a subgroup of index 2 in G, so $|G| = 2v$. Observe also that $G \simeq D_{2v}$.

3.1.9 Example (Symmetry groups of regular solids). We saw above (Corollary 2.2.8) that the orientation-preserving halves of the symmetry groups of the regular solids are finite Kleinian groups.

Remark. One may show (see [34, Theorem C.10]) that the nontrivial finite Kleinian groups may be classified as follows: such a group is cyclic, dihedral, or conjugate in $\mathbb{M}$ to the symmetry group of a regular solid.

3.1.10 Example (Bianchi groups). Suppose R is a discrete subring of $\mathbb{C}$. Then $\mathrm{PSL}(2, R)$ is a discrete subgroup of $\mathrm{PSL}(2, \mathbb{C})$: indeed, if $A_n \to I$ then each component tends to a value in R, contradicting the lack of limit points. Let $\mathcal{O}_d$ for $d \in \mathbb{Z}_{>0}$ squarefree be the ring of integers of the quadratic imaginary number field $\mathbb{Q}(\sqrt{-d})$ (recall, the ring of an integers of a number field F is the integral closure of $\mathbb{Z}$ in F). The groups $\mathrm{PSL}(2, \mathcal{O}_d)$ are the **Bianchi groups**.

Note that O_d is generated by the two elements 1 and ζ_d, where

$$\zeta_d = \begin{cases} \sqrt{-d} & d \equiv \pm 1 \pmod 4, \\ \frac{1+\sqrt{-d}}{2} & d \equiv 2 \pmod 4. \end{cases}$$

Hence the translations $z \mapsto z + 1$ and $z \mapsto z + \zeta_d$ are elements of $\mathrm{PSL}(2, O_d)$, and so Bianchi groups always have a cusp at ∞.

3.1.11 Example (Coxeter groups). If $P \subseteq H^3$ is a convex acute-angled polyhedron, with dihedral angles submultiples of π, the Coxeter group generated by reflections in the faces of P is discrete; the index-2 subgroup of orientation-preserving elements is therefore a Kleinian group.

3.1.12 Example (Modular groups). The **modular groups** are $\mathrm{SL}(2, \mathbb{Z})$ and its subgroups of finite index; see [36, Chapter 4]. We have that $\mathrm{SL}(2, \mathbb{Z})$ is generated by two elements,

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix};$$

further:

- The elliptic points of $\mathrm{SL}(2, \mathbb{Z})$ in $\hat{\mathbb{C}}$ are the orbits of i and $\exp(\pi i/3)$;
- The cusps of $\mathrm{SL}(2, \mathbb{Z})$ in $\hat{\mathbb{C}}$ are $\mathbb{Q} \cup \{\infty\}$.

3.1.13 Example. Define $X, Y_4 \in \mathrm{SL}(2, \mathbb{C})$ by

$$X = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad Y_4 = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix},$$

and let $G = \langle X, Y_4\rangle$. The isometric circles of Y_4 and Y_4^{-1} are $|4z + 1| = 1$ and $|4z - 1| = 1$. Consider the set

$$F \coloneqq \{z = x + yi \in \mathbb{C} : -1/2 < x < 1/2, |4z + 1| > 1, |4z - 1| > 1\} \cup \{\infty\}$$

depicted in Fig. 3.1. The cusp points of G are (the translates of) $0, \infty, 1/2, -1/2$ (corresponding to the elliptic elements Y_4, X, XY_4^{-1}, and $(XY_4^{-1})^{-1}$); notice these correspond to the 'cusps' visible on the diagram of F.

One may also observe also that F has the following properties (where $\Omega = \Omega(G)$):

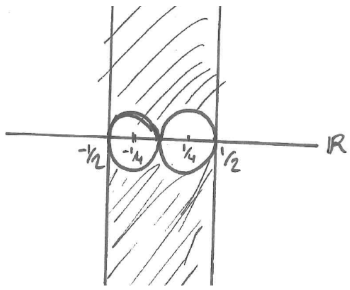


Figure 3.1: The set F of Example 3.1.13.

1. F is an open subset of Ω;
2. The members of $\{gF : g \in G\}$ are mutually disjoint;
3. For every $z \in \Omega$, there is some $g \in G$ with $g(z) \in \overline{F}$;
4. The intersection $\partial F \cap \Omega$ consists of a countable (in fact, finite) number of curves, and for each such curve s there is another such curve s' (not necessarily distinct from s) such that $g(s) = s'$ for some $s \in S'$.

We shall prove most of these statements in the following sections (we will see that F is a **fundamental domain** for G).

3.2 Some more results on isometric circles

Fix some $G \leq \mathbb{M}$ be Kleinian; in this section we will prove some more results about the isometric circle of elements $g \in G$, as preparation for our study of limit points. Recall that the isometric circle of g is the unique circle $S \subseteq \hat{\mathbb{C}}$ such that $g\restriction_S : S \to g(S)$ is an isometry.

We now collect some easy results, and some other results which follow directly from earlier results. Note that the factorisation result is proved as part of the proof of Corollary 1.4.3.

3.2.1 Lemma. *If* $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2,\mathbb{C})$ *is a representative for $g \in G$, then the isometric circle S of g is the set*

$$|cz + d| = 1,$$

with centre $g^{-1}(\infty) = -d/c$ and radius $\left|c^{-1}\right|$. Further, g factorises as $g = s \circ r \circ q$, where q is circle inversion in S and both r and s are Euclidean motions (distance-preserving affine maps).

Let $\mu(X)$ denote the spherical area of a set $X \subseteq \hat{\mathbb{C}}$ (defined similarly to the chordal metric, via the pullback of the spherical area on S^2 via stereographic projection: compare Eq. (1.2.9) and the preceeding discussion).

3.2.2 Lemma. *Let U be a nice neighbourhood of some $z \in {}^\circ\Omega$. Then*

$$\sum_{g \in G} \mu(gU) < \infty.$$

Proof. The gU are all disjoint, and $\mu(S^2) < \infty$.

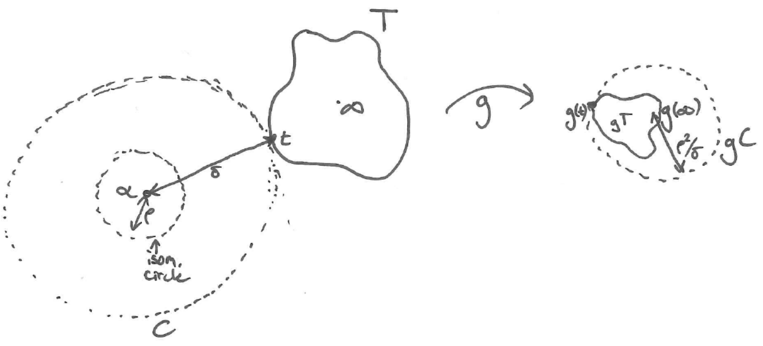


Figure 3.2: Figure for Lemma 3.2.3.

Remark. From this lemma, we obtain another proof of Lemma 3.1.3: it is well-known that if $(x_a)_{a\in A}$ is a family of nonnegative real numbers, then $\sum_{a\in A} x_a < \infty$ only if $x_a = 0$ for all but countably many a; in the lemma, all the $\mu(gU)$ are positive and so the index set G is countable.

For the following lemma and theorem, let $\operatorname{diam} S$ for $S \subseteq \hat{\mathbb{C}}$ denote the Euclidean diameter (i.e. $\sup_{x,y\in S} |x-y|$).

3.2.3 Lemma. *Let $g \in \mathbb{M}$ not fix ∞, and let T be a closed set not containing $\alpha := g^{-1}(\infty)$. Let $\delta = d(\alpha, T)$, and let ρ be the radius of the isometric circle of g. Then*

$$\operatorname{diam}(gT) \le 2\rho^2/\delta,$$

and if $\infty \in T$ then

$$\rho^2/\delta \le \operatorname{diam}(gT).$$

Proof. See Fig. 3.2. Since T is closed, $d(\alpha, T) > 0$. Now note that T lies outside the (open) ball centred at α of radius δ; the circle $C := S(\alpha, \delta)$ is sent by g to a circle gC of radius ρ^2/δ, since the only component of g which changes the radius of circles is the reflection across the isometric circle, which is centred at α (Lemma 3.2.1). Note now that g sends the exterior of C (i.e. the connected component of $\hat{\mathbb{C}} \setminus C$ containing ∞) to the interior of the circle gC (the connected component of $\hat{\mathbb{C}} \setminus gC$ not containing ∞). In particular, $gT \subseteq gC$, so $\operatorname{diam}(gT) \le \operatorname{diam}(gC) = 2\rho^2/\delta$.

If $\infty \in T$ then $g(\infty)$ (the centre of gC) lies in gT; on the other hand, since T is closed there is a point $t \in T$ such that $d(t, \alpha) = \delta$ and so $gt \in gC$; hence $\operatorname{diam}(gT) \ge |gt - g(\infty)| = \rho^2/\delta$.

3.2.4 Theorem. *If $\infty \in {}^\circ\Omega$, then*

$$\sum_{g\in G\setminus 1} |c|^{-4} < \infty$$

(where c denotes the bottom-left entry of an arbitrary matrix representative of g).

Proof. See Fig. 3.3. Let U be a nice neighbourhood of ∞; without loss of generality, $U = S(\infty, \rho)$. Suppose $g \in G$ is nontrivial; since $\infty \in {}^\circ\Omega$, $g(\infty) \neq \infty$ and so g has an isometric circle C, say $S(\alpha, |c|^{-1})$. Since U is a nice neighbourhood, $\alpha \notin U$ (otherwise $g(\alpha) = \infty$, so $\infty \in gU \cap U$). By increasing ρ to ρ' if necessary, we may assume $\delta := d(\alpha, U) > 0$. Note that $\delta \le \rho \le \rho'$; hence by applying Lemma 3.2.3 we have

$$\operatorname{diam}(gU) \ge |c|^{-2}/\delta \ge |c|^{-2}/\rho' \ge |c|^{-2}/\rho.$$

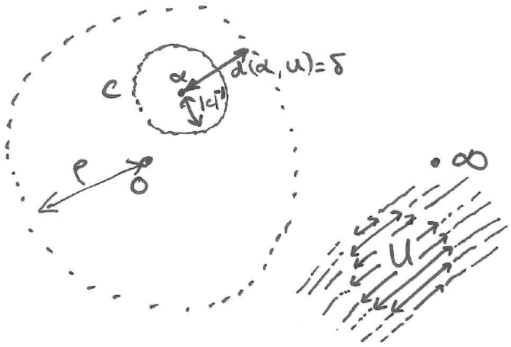


Figure 3.3: Figure for Theorem 3.2.4.

Since gU is a circular disc contained in $\hat{\mathbb{C}} \setminus U$ (in particular, gU is bounded) and since stereographic projection has bounded distortion there exists some $K \in \mathbb{R}_{>0}$ such that $\mu(gU) \geq K^{-1}\,\mathrm{diam}(gU)^2$. In fact, K is dependent only on the bound for gU, and this bound — ρ — is independent of g. Observe also that ρ is independent of g even though ρ' was not. Thus:

$$\sum_{g \in G \setminus 1} |c|^{-4} \leq \sum_{g \in G \setminus 1} \rho^2\,\mathrm{diam}(gU)^2 \leq \rho^2 K \sum_{g \in G \setminus 1} \mu(gU);$$

and we may apply Lemma 3.2.2 to complete the proof.

3.2.5 Corollary. *Let (g_n) be a sequence of distinct elements of G, with $\infty \in {}^\circ\Omega$; for each n, let r_n be the radius of the isometric circle of g_n. Then $r_n \to 0$.*

Proof. If g_n has matrix representative with bottom-left entry c_n, then note that $r_n = |c_n|^{-1}$ for all n; by Theorem 3.2.4, $|c_n|^{-4} \to 0$ as $n \to \infty$ and so $|c_n|^{-1} \to 0$ as well.

3.2.6 Proposition. *Let (g_n) be a sequence of elements of G, and let $x, y \in \hat{\mathbb{C}}$ be such that $g_n(\infty) \to x$ and $g_n^{-1}(\infty) \to y$. Let $C_n = S(w_n, r_n)$ be the isometric circle of g_n, and let $C_n' = S(w_n', r_n')$ be the isometric circle of g_n'. Then*

$$w_n \to x, \qquad w_n' \to y, \qquad r_n = r_n' \to 0.$$

Proof. Let $A_n = \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix} \in \mathrm{SL}(2, \mathbb{C})$ be a representative for g_n, so g_n^{-1} is represented by $A_n^{-1} = \begin{bmatrix} d_n & -b_n \\ -c_n & a_n \end{bmatrix}$; we have that $g_n(\infty) = a_n/c_n \to x$ and that $g_n^{-1}(\infty) = -d_n/c_n \to y$. Now note that C_n is the locus of $z \in \hat{\mathbb{C}}$ such that $|c_n z + d_n| = 1$, i.e. the centre of C_n is the point w_n with $c_n w_n + d_n = 0$ and then $w_n = -d_n/c_n \to y$; similarly, C_n' has centre the point w_n' with $-c_n w_n' + a_n = 0$, i.e. $w_n' = a_n/c_n \to x$. Finally, the statement about the radii is a direct application of Corollary 3.2.5.

3.3 Limit points

Let $G \leq \mathbb{M}$ be Kleinian. For fixed $w \in \hat{\mathbb{C}}$, write $\Lambda(w)$ for the set of points $z \in \hat{\mathbb{C}}$ such that there exists a sequence (g_n) of distinct elements of G with $g_n w \to z$. (Note, we do not require the $g_n w$ to

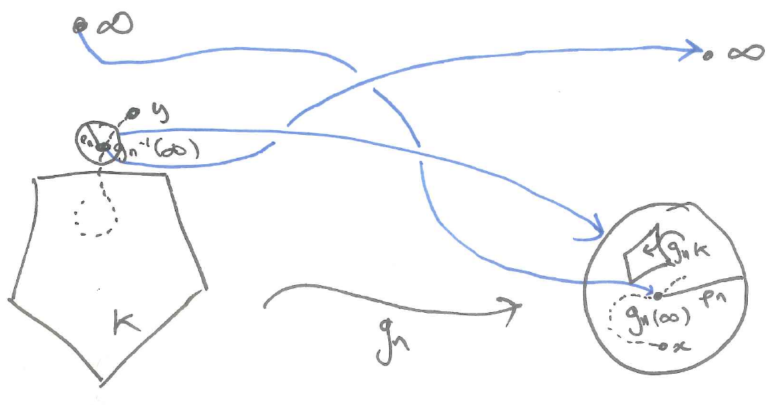


Figure 3.4: The proof of the compact convergence lemma, Lemma 3.3.2.

be distinct.) A **limit point** of G is a point which lies in $\Lambda(w)$ for some $w \in {}^\circ\Omega$; we denote by Λ the set $\bigcup_{w \in {}^\circ\Omega} \Lambda(w)$; this is called the **limit set** of G.

3.3.1 Definition. Let X be a topological space, and (Y, d) a metric space. A sequence of functions $f_n : X \to Y$ ($n \in \mathbb{N}$) is said to **converge uniformly on compact subsets** to some $f : X \to Y$ if for all $K \subseteq Y$ compact,

$$\lim_{n\to\infty} \sup_{k \in K} d(f_n(k), f(k)) = 0.$$

The following lemma is fundamental.

3.3.2 Lemma. *1. Let $x \in \Lambda$; then there exists $y \in \Lambda$, and a sequence (g_n) of distinct elements of G, such that the maps $g_n : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ converge uniformly to the constant function x on compact subsets of $\hat{\mathbb{C}} \setminus \{y\}$.*

2. Further, if (g_n) is an arbitrary sequence of distinct elements of G then there is a subsequence (g_m) and $x, y \in \Lambda$ such that $g_m \to x$ uniformly on compact subsets of $\hat{\mathbb{C}} \setminus \{y\}$.

Proof. Since x is a limit point, we may find a sequence $(g_m) \subseteq G$ of distinct elements and an element $z_0 \in {}^\circ\Omega$ such that $g_m z_0 \to x$. By conjugation, we may assume $z_0 = \infty$. Since $\hat{\mathbb{C}}$ is compact, there is a subsequence (g_n) of (g_m) such that the sequence $g_n^{-1}(\infty)$ converges to some point y; by construction, $y \in \Lambda$. By Proposition 3.2.6, we have that the centres of the isometric circles of g_n tend to x, the centres of the isometric circles of g_n^{-1} tend to y, and the isometric circle radii tend to 0. Let $K \subseteq \hat{\mathbb{C}} \setminus \{y\}$ be compact; let N be sufficiently large that for all $n > N$, $B(g_n^{-1}(\infty), \rho_n)$ lies outside K (Fig. 3.4). Note that for such n, the exterior of the isometric circle of g_n is mapped to the interior of the isometric circle of g_n^{-1}; hence gK lies in the interior of the isometric circle of g_n^{-1}, which is $B(g_n(\infty), \rho_n)$; thus as $n \to \infty$, $gK \to \lim_{n\to\infty} g_n(\infty) = x$ uniformly.

For the second part of the lemma, choose (again by compactness) a subsequence (g_m) such that $g_m(\infty)$ converges to some $x \in \hat{\mathbb{C}}$ and such that $g_m^{-1}(\infty)$ converges to some $y \in \hat{\mathbb{C}}$. By suitable conjugations, we may assume $\infty \in {}^\circ(G)$ and so $x, y \in \Lambda$. By a similar argument to the first part, we see that for $K \subseteq \hat{\mathbb{C}} \setminus \{y\}$ compact we have $gK \to \lim_{n\to\infty} g_n(\infty) = x$ uniformly.

3.3.3 Theorem. *Let G be Kleinian of the second type. The set Λ may alternatively be characterised as:*

1. The set $\Lambda(z_0)$ of limit points, for any fixed $z_0 \in {}^\circ\Omega$;

2. *The set $\overline{\Lambda_0}$, where Λ_0 is the set of non-elliptic points of G;*

3. *The set $\hat{\mathbb{C}} \setminus \Omega$.*

*We call this set the **limit set** of G.*

Proof. 1. Fix $z_0 \in {}^\circ\Omega$ arbitrary; clearly $\Lambda(z_0) \subseteq \Lambda$. Suppose on the other hand that $w \in \Lambda$; then there is some $z \in {}^\circ\Omega$ and some sequence (g_n) of distinct elements of G such that $g_n(z) \to w$; by the proof of Lemma 3.3.2, there is a subsequence of the g_n which converges uniformly to the constant function w on compact subsets; in particular, the value of z_0 under this subsequence tends to w and so $w \in \Lambda(z_0)$.

2. This is harder, we shall prove it later as Corollary 3.3.22.

3. First, note that if $x \in \Lambda$ then every neighbourhood of x has infinitely many translates of some point, so $x \notin \Omega$. Thus $\Lambda \cap \Omega = \emptyset$.

   Assume now that $x \notin \Omega$; we show that $x \in \Lambda$. For every neighbourhood $U \ni x$, there are infinitely many $g \in G$ with $gU \cap U \neq \emptyset$. Hence by taking the neighbourhoods $B(x, 1/m)$ for $m \in \mathbb{N}$ we can find a sequence (g_m) of distinct elements of G and a sequence of points $z_m \in \hat{\mathbb{C}}$ such that $g_m z_m \to x$ and $z_m \to x$. By the proof of Lemma 3.3.2, we can find a subsequence (g_n) and limit points w, y such that $g_n \to w$ uniformly on $\hat{\mathbb{C}} \setminus \{y\}$. If $x = y$, then $x \in \Lambda$. If $x \neq y$, then the points z_m do not accumulate at y, so we may use the convergence property away from y to see $g_n z_n \to w$, thus $x = w \in \Lambda$.

Remark. Compare (1) of the above theorem with [6, Theorem 5.3.9]: if we assume G is non-elementary (see below) we may remove the requirement for z_0 to be a point of free discontinuity.

3.3.4 Theorem. *If G is Kleinian of the second type, then the set $\Lambda(G)$ is closed, G-invariant, and nowhere dense in $\hat{\mathbb{C}}$. If G is Kleinian of the first type then $\Lambda(G) = \hat{\mathbb{C}}$ so $\Lambda(G)$ is closed, G-invariant, and dense in $\hat{\mathbb{C}}$.*

Proof. The statements about Kleinian groups of the first type are trivial, so let G be Kleinian of the second type.

Closedness. Let (x_n) be a sequence of points in Λ, with $x_n \to x \in \hat{\mathbb{C}}$. By Lemma 3.3.2, or equivalently the equality $\Lambda = \Lambda(z_0)$ of the above theorem, there exists a single point $z_0 \in \hat{\mathbb{C}}$ and sequences of distinct elements $g_{m,n}$ of G such that $\lim_{m\to\infty} g_{m,n}(z_0) = x_n$ for all n. Without loss of generality, the points $\{x_n\}$ and x are all distinct (otherwise, x_n is eventually equal to x and so $x \in \Lambda$). For all n, let δ_n be the minimal distance from x_n to the other elements of the set $\{x_m\}$. For all n, pick $k(n)$ such that $d(g_{k(n),n}(z_0), x_n) < \delta_n/2$ (Fig. 3.5); the effect of this constraint is to make the $g_{k(n),n}$ all distinct. Now note that $\lim_{n\to\infty} g_{k(n),n}(z_0) = x$, since for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $d(x_n, g_{k(n),n}(z_0)) < \varepsilon/2$ and $d(x_n, x) < \varepsilon/2$ for all $n > N$.

G-invariantness. Let $x \in \Lambda$, $g \in G$. Pick $z \in \hat{\mathbb{C}}$ and (g_n) distinct in G such that $g_n z \to x$; then $g \circ g_n z \to gx$, as left-multiplication by g is continuous.

Nowhere dense. Recall, Y is nowhere dense in a topological space X if every non-empty open subset V of X contains a non-empty open subset U of X with $V \cap Y = \emptyset$. Here, if $V \subseteq \hat{\mathbb{C}}$ is non-empty then either $V \cap \Lambda = \emptyset$ (so we may take $U = V$), or $V \cap \Lambda \neq \emptyset$, in which case V contains points of ${}^\circ\Omega$ (by definition of Λ) and so $U := {}^\circ\Omega \cap V$ is non-empty and open, with the property that $U \cap \Lambda = \emptyset$ (since $\Lambda = \hat{\mathbb{C}} \setminus \Omega$ and ${}^\circ\Omega \subseteq \Omega$).

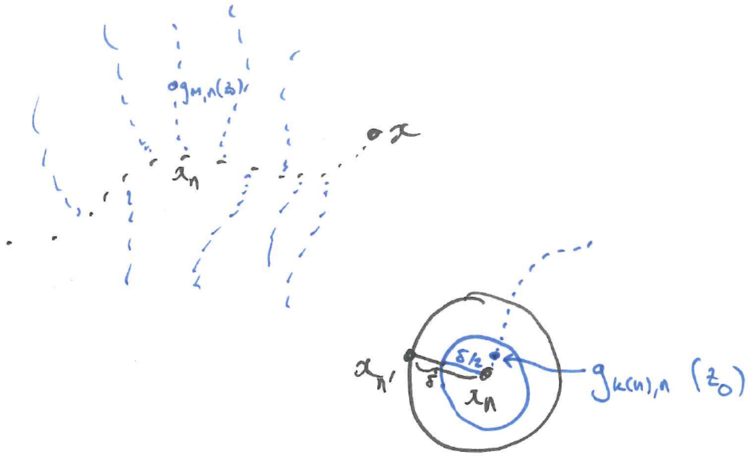


Figure 3.5: The sequential limit of a sequence of limit points.

Recall that a set S is perfect if every point of S is a limit point of S. It is standard that such sets are uncountable.

3.3.5 Corollary. *If Λ contains more two points, then Λ is perfect.*

Proof. Suppose $|\Lambda| \geq 3$. Let $x \in \Lambda$, then by Lemma 3.3.2 there exists a sequence (g_n) of distinct elements of G and some $y \in \Lambda$ with $g_n(z) \to x$ for all $z \neq y$. In particular, we may pick $x_1, x_2 \in \Lambda$ such that $\{y, x_1, x_2\}$ are all distinct and such that $g_n(x_1) \to x$ and $g_n(x_2) \to x$. If $g_m(x_1) = x = g_m(x_2)$ for some m, then $x_1 = x_2$; thus for all m, either $g_m(x_1)$ or $g_m(x_2)$ is distinct from x. Thus at least one of the sequences $g_m(x_i)$ has an infinite subsequence of distinct limit points of Λ (here we use G-invariantness) tending to x.

We say that G is **elementary** if the action of G on $\overline{H^3}$ has a finite orbit; otherwise, we say G is **non-elementary**. If G is elementary, then we variously say that:

- G is **of elliptic type** if G has a finite orbit in H^3;
- G is **of parabolic type** if G has a fixed point in ∂H^3 and no other finite orbits in $\overline{H^3}$;
- G is **of loxodromic type** if it is neither of elliptic type nor of parabolic type.

Compare with Proposition 2.3.13.

3.3.6 Lemma. *If G is elementary Kleinian of elliptic type, then G fixes a point in H^3.*

Proof. Let $\{x_1, \dots, x_n\}$ be a finite orbit of G in H^3. If $g \in G$, then the powers $g^m(x_1)$ for $m \in \mathbb{N}$ cannot be all distinct, so there is some m with $g^m(x_1) = x_1$. In particular, g^m has a fixed point in H^3, so g^m is elliptic. Now observe that if g^m is elliptic then g is elliptic: indeed, g^m is conjugate to $z \mapsto k^2 z$ with $|k^2| = 1$, so g is conjugate to $z \mapsto k^2/mz$ where $|k^2/m| = 1^{1/m} = 1$. Thus each element of G is elliptic and by Theorem 2.4.7 G has a fixed point in H^3.

3.3.7 Lemma. *Discrete subgroups of $O(3)$ are finite.*

Proof. It suffices to note that $O(3)$ is compact: if $A \in O(3)$ then the columns of A are orthonormal and so $\|A\|^2 = 1 + 1 + 1 = 3$ (so $O(3)$ is bounded), and $AA^t = 1$, so A is cut out by polynomials in its entries and is closed. ∎

3.3.8 Lemma. *A Kleinian group is finite iff every element has finite order.*

Proof. Every element of a finite group has finite order. Conversely, let $G \leq \mathbb{M}$ be Kleinian with every element of finite order. By Theorem 2.4.7, all the elements of G have a common fixed point; conjugating this fixed point to 0, we see that G is conjugate to a discrete subgroup of $O(3)$ and hence is finite by Lemma 3.3.7. ∎

3.3.9 Proposition. *If $f, g \in \mathbb{M}$ are nontrivial, where f is loxodromic and f, g have exactly one shared fixed point, then $\langle f, g\rangle$ is not discrete.*

Proof. Assume the common fixed point is ∞ and conjugate f to the transformation $f(z) = k^2 z$ for $k^2 \in \mathbb{C}$; if ∞ is the repelling fixed point of f, replace f with f^{-1}. Then g is of the form $g(z) = az + b$, and since $g(0) \neq 0$ we have $b \neq 0$. Observe that $f^{-n} g f^n(z) = az + k^{-2n} b$. As $\left|k^2\right| > 1$, we see that $\left|g^{-n} f g^n\right|^2 = |a|^2 + \left|k^{-2n} b\right|^2 \to 0$ as $n \to \infty$, so the sequence of distinct elements $(g^{-n} f g^n)$ has a limit point. ∎

3.3.10 Lemma. *Let G contain two parabolic elements with distinct fixed points. Then G contains a loxodromic element.*

Proof. Suppose that $f, g \in G$ are parabolic with distinct fixed points; by conjugating appropriately, we have normal forms

$$A = \begin{bmatrix} 1 & p \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ q & 1 \end{bmatrix}$$

for f and g respectively with p, q nonzero; now note that

$$AB = \begin{bmatrix} 1 + pq & p \\ q & 1 \end{bmatrix}, \qquad AB^{-1} = \begin{bmatrix} 1 - pq & p \\ -q & 1 \end{bmatrix};$$

since fg is not loxodromic, it is either parabolic or elliptic; note that fg cannot fix ∞, otherwise f, g would have a shared fixed point, and so if fg were parabolic with fixed point x there would be $r \neq 0$ with $1 - pq = 1 - rx$ and $1 = 1 + rx$; this occurs only if $x = 0$ and $pq = 0$ which is not allowed by assumption on p and q. Hence fg must be elliptic, with $\operatorname{tr}^2 fg = (2 + pq)^2 \in [0, 4)$; thus $2 + pq \in (-2, 2)$ and $pq \in (-4, -2)$, so $2 - pq \in (2, 4)$ so $\operatorname{tr}^2 fg^{-1} \in (4, 16)$ — and fg^{-1} is loxodromic (contradiction). ∎

3.3.11 Lemma. *If G is elementary of loxodromic type, then G leaves invariant a unique hyperbolic line in H^3.*

Proof. Since G is not of elliptic type, the finite orbits of G are all contained in ∂H^3. Let $\{u_1, \dots, u_n\}$ be the union of a finite number of finite orbits of G. By the orbit-stabiliser theorem, each stabiliser $\operatorname{Stab}_G u_i$ is of finite index in G. Thus $H = \operatorname{Stab}_G u_1 \cap \cdots \cap \operatorname{Stab}_G u_n$ is of finite index in G and fixes each u_i. If $n \geq 3$, then H (hence G) is of elliptic type: this is because each element of H must fix three elements and so must have a fixed point in H^3, so every element is elliptic. Hence $n = 1$ or $n = 2$. If $n = 1$, then either G is of elliptic type or G is of parabolic type. Thus $n = 2$; it is clear that G leaves invariant the line $[u_1, u_2]$. ∎

3.3.12 Theorem. *The following are equivalent:*

1. *G is elementary;*

2. $|\Lambda(G)| \leq 2$ *(and further,* $|\Lambda(G)|$ *is* $0, 1, 2$ *as G is of elliptic, parabolic, or loxodromic type respectively);*

3. $|\Lambda(G)| < \infty$.

Proof. The implication $|\Lambda(G)| \leq 2 \implies |\Lambda(G)| \leq \infty$ is trivial. Now suppose $|\Lambda(G)| \leq \infty$. Since $\Lambda(G)$ is G-invariant, $\Lambda(G)$ is a union of G-orbits; it is immediate that each such orbit is finite.

Suppose G is elementary. We have three cases.

1. G is of elliptic type. By Lemma 3.3.6, G has a fixed point z in H^3, in particular each element of G is elliptic, and by discreteness each such element has finite order (otherwise, pick a sequence of distinct elements of the group, the multipliers of the elements lie on S^1 and so there is a subsequence of elements with multipliers converging to a fixed value, contradicting discreteness since there is exactly one element with a given multiplier); in particular, by Lemma 3.3.8, G is finite and so cannot have any limit points.

2. G is of parabolic type. Let z be the fixed point of G in $\hat{\mathbb{C}}$; by conjugation, assume $z = \infty$. Observe that G must contain a parabolic or loxodromic element, otherwise every nontrivial element is elliptic and thus by Theorem 2.4.7 G fixes a point in H^3.

   If g contains a parabolic element, then every element of G fixes ∞ (all other orbits are infinite) and so existence of a loxodromic element $f \in G$ contradicts discreteness by Proposition 3.3.9. Hence every element of G is parabolic or elliptic, and there is a unique limit point (namely, ∞) — to see this, use discreteness to see that each elliptic element is of finite order (as in the first case above) so no finite limit point comes from adding in the elliptic elements.

   On the other hand suppose G consists entirely of loxodromic and elliptic elements; then if $g \in G$ is an arbitrary loxodromic element, with fixed points 0 and ∞ (by conjugation), every other element must leave $\{0, \infty\}$ invariant (again, by Proposition 3.3.9). Since ∞ is a shared fixed point, this implies that 0 is a shared fixed point, contradicting uniqueness of the fixed point for G; thus this case cannot occur and if G is of parabolic type it contains only parabolic and elliptic elements.

3. G is of loxodromic type. By the arguments above, we see that G cannot be entirely elliptic (otherwise G fixes a point in H^3, so has a finite orbit in H^3, so is elliptic type). Suppose for contradiction that G contains no loxodromic element. Then there is some $g \in G$ parabolic, say with fixed point ∞. If every parabolic element of G had fixed point ∞, we see easily that G would be of parabolic type. Thus there is some parabolic $h \in G$ which does not fix ∞; by Lemma 3.3.10 this implies existence of a loxodromic element, contradiction. We conclude that G contains a loxodromic element g, with fixed points $w, z \in \hat{\mathbb{C}}$, say with w attracting and z repelling. Note that for fixed $z_0 \in {}^\circ\Omega$ we have $g^n z_0 \to w$ and $g^{-n} z_0 \to z$, so $w, z \in \Lambda$. By Lemma 3.3.11, G leaves the unique line $[w, z]$ invariant in H^3. Suppose x is a limit point; then x is the limit of a sequence $g_n z_0$ for (g_n) distinct elements of G. Note that the g_n are eventually not loxodromic, for the images of loxodromic elements must tend to w or z. But G is discrete, so we must have that x is the fixed point of infinitely many parabolic elements; but note, if $h \in G$ is parabolic then gh is loxodromic whenever h is loxodromic, and gh has different fixed points to h, contradiction.

This shows that G elementary implies that $|\Lambda(G)| \leq 2$. ☡

There is in fact a fourth equivalent definition of elementariness:

3.3.13 Proposition. *A Kleinian group G is elementary iff it contains an abelian subgroup G' of finite index.*

Proof. Suppose G is an elementary Kleinian group.

- If G is elliptic, then G is finite (by case 1 in the proof of Theorem 3.3.12) and hence we may take $G' = G$.
- If G is loxodromic, then by conjugation we may assume that 0 and ∞ are the two limit points of G (and that these are the only fixed points of any element of G in $\overline{\mathbb{H}^3}$); so $\{0, \infty\}$ is left invariant by G. Let G_0 be the subgroup of G fixing 0. By the Orbit-Stabiliser theorem, $[G : G_0] \le 2$. Each element of G_0 is a Poincaré extension of an element of $\lambda O(2)$ for some $\lambda \in \mathbb{R}_{>0}$ (by Corollary 1.3.9). Let $\rho : G_0 \to \mathbb{R}_{>0}$ be the map sending $\lambda A \mapsto \lambda$; the kernel of this map is $G_0 \cap O(2)$, which is finite by discreteness (since $O(2)$ is compact). The orbit $G_0 j$ is discrete (because it cannot accumulate anywhere!) and so $\rho(G_0)$ is an infinite discrete subgroup of $\mathbb{R}_{>0}$. Hence there is some $s > 1$ generating $\rho(G_0)$: that is, $\rho(G_0) = \{s^m : m \in \mathbb{Z}\}$. Thus, ρ exhibits G_0 as the extension of an infinite cyclic (hence abelian) group by a finite group and hence G is a finite extension of an abelian group G' as desired.
- If G is parabolic, we shall use the following fact:

  Claim. *Let G be a group of isometries of $\mathbb{R}^n$. Then G has an abelian normal subgroup N of finite index containing all the translations in G, and the index of N in G is bounded by a number depending only on n.*

  One can choose N to be the subgroup generated by all of the elements $\phi = a + A$ in G (where $a \in \mathbb{R}^n$ and $A \in G$) such that $\|A - 1\| < 1/2$; the proof is given in detail as [38, lemma 7 of section 5.4].

  But, of course, if G is parabolic then we may assume the global fixed point is ∞, and so G acts as a discrete group of isometries on $\mathbb{R}^3$. Then the claim immediately gives the result.

This proves one direction.

Suppose now that G is a Kleinian group, and that $G' \le G$ is a finite index abelian subgroup. Observe that G' is elementary, by Theorem 2.4.3. Let $x \in \overline{H^3}$ such that $G'x$ is finite. Since $[G : G'] < \infty$, there exist coset representatives $\gamma_1, ..., \gamma_m \in G$ such that $G = \bigcup_{i=1}^m \gamma_i G'$ and so $Gx = \bigcup_{i=1}^m \gamma_i G' x$, which is finite. This proves that G is elementary. ∎

We now perform a study of non-elementary groups.

3.3.14 Lemma. *If $g \in \mathbb{M}$ is elliptic or parabolic, and if g does not fix ∞, then the isometric spheres of g and g^{-1} intersect.*

Proof. Suppose g is elliptic; conjugate such that g fixes 0 and 1; then we have matrix representatives

$$\begin{bmatrix} \overline{k} & 0 \\ \overline{k} - k & k \end{bmatrix} \qquad \begin{bmatrix} k & 0 \\ k - \overline{k} & \overline{k} \end{bmatrix}$$

for g and g^{-1} respectively; now note, the isometric circle of g has equation $\left|(\overline{k} - k)z + k\right| = 1$, and that for g^{-1} has equation $\left|(k - \overline{k})z + \overline{k}\right| = 1$. Hence both have radius $r = \left\|k - \overline{k}\right\|^{-1}$, and the centres are respectively $-\overline{k}/(k - \overline{k})$ and $-k/(\overline{k} - k)$, with distance

$$\left|-\overline{k}/(k - \overline{k}) - \left(-k/(\overline{k} - k)\right)\right| = \left|-\overline{k}/(k - \overline{k}) - k/(k - \overline{k})\right| = r\left|-\overline{k} - k\right| = r\left|\overline{k} + k\right|;$$

note, $\overline{k} + k = 2\operatorname{Re}\overline{k} < 2$. The isometric circles, of radius r, therefore have centres a distance apart less than $2r$ and so must intersect.

A similar but easier computation shows the parabolic case.

3.3.15 Lemma. *If G has no loxodromic elements, contains a parabolic element, and all the parabolic elements have a common fixed point, then G is conjugate to a group of Euclidean motions, and G has a unique limit point. In particular, G is elementary.*

Proof. Conjugate the common fixed point of the parabolic elements to ∞. It suffices to show that all the elliptic elements of G then have a fixed point at ∞; suppose for the sake of contradiction that there is some $f \in G$ elliptic with finite fixed points, and without loss of generality conjugate the fixed points of f to 0 and 1. Let g be an arbitrary parabolic element. We may (by the normal form results) choose matrices A and B for f and g respectively of the form

$$A = \begin{bmatrix} \overline{k} & 0 \\ \overline{k} - k & k \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & p \\ 0 & 1 \end{bmatrix}, \qquad \text{so } B^n A = \begin{bmatrix} \overline{k} + p^n(\overline{k} - k) & p^n k \\ \overline{k} - k & k \end{bmatrix}$$

and hence the isometric circle of $g^n f$ is exactly that of f for all $n \in \mathbb{N}$. Similarly, we have

$$A^{-1}B^{-n} = \begin{bmatrix} k & -p^n k \\ k - \overline{k} & \overline{k} + p^n(\overline{k} - k) \end{bmatrix},$$

so the isometric circle of $f^{-1}g^{-n}$ has the same radius as that of f but centre $-(\overline{k} + p^n(\overline{k} - k))/(k - \overline{k}) = p^n - \overline{k}/(k - \overline{k})$, i.e. the isometric circle of fg^n is the image of the isometric circle of f^{-1} under a translation by p^n.

Let S be the isometric circle of f, and S' the isometric circle of f^{-1}. Observe now that the cyclic group generated by f acts discontinuously on $\hat{\mathbb{C}}$ and so there is some $m > 0$ such that S and $f^m S'$ are disjoint; but by the observations above, this shows that the isometric circles of $g^m f$ and $f^{-1}g^{-m}$ are disjoint; note that these elements are not loxodromic by assumption, and so by Lemma 3.3.14 we obtain a contradiction.

It is now easy to see that the only limit point of a discrete group of Euclidean motions G with a parabolic element is the point ∞.

3.3.16 Proposition. *If G is a non-elementary group, then G has a loxodromic point.*

Proof. By Lemma 3.3.8, if every nontrivial element of G was elliptic then G would be finite, in particular G would have no limit points, so would be elementary by Theorem 3.3.12. Hence G contains a nontrivial element which is parabolic or loxodromic.

If G contains no loxodromic elements and all the parabolic elements have a common fixed point, then by Lemma 3.3.15 G is conjugate to a group of Euclidean motions and is elementary. Hence if G contains no loxodromic elements, there must be two parabolic elements $f, g \in G$ with distinct fixed points; but this contradicts Lemma 3.3.10.

3.3.17 Corollary. *If G is finite, then G is elementary.*

Proof. By Corollary 2.4.8, a finite Kleinian group consists entirely of elliptic elements.

We now obtain the following lemma, which is fundamental in the non-elementary case as it strengthens Lemma 3.3.2. We see that the 'generic case' of a non-elementary group is loxodromic.

3.3.18 Proposition. *If G is non-elementary and $x \in \Lambda$, then the G-translates of x are dense in Λ.*

Proof. By Proposition 3.3.16, G contains a loxodromic element f. If x is not fixed by f, then there is some element of G which does not fix x. Otherwise, if x is fixed by f, then take a limit point y not fixed by f; by Lemma 3.3.2 there exists a fixed point z of f and a sequence of elements (g_n) of G with $g_n z \to y$. For every $g \in G$, the fixed point sets of f and gfg^{-1} are either identical or disjoint by Proposition 3.3.9. This means that eventually $g_n f g_n^{-1}$ cannot have a fixed point at z, since $z \neq y$. In particular, there exists an element $g \in G$ with the fixed point sets of f and gfg^{-1} disjoint! In particular, x is not a fixed point of gfg^{-1}, and so there is some element of G which does not fix x.

Overall, we have some $g \in G$ such that $gx \neq x$. Let $z \in \Lambda$ be arbitrary; by Lemma 3.3.2, some subsequence of the sequence (g^n) has the property $g^n(x) \to z$ which completes the proof.

Remark. Observe that what goes wrong in the elementary case is that there might be some limit point not moved by any element of G. (For instance, take the group generated by the translation $z \mapsto z+1$.)

3.3.19 Corollary. *If G is non-elementary, then G is of the second kind iff there exists some $z \in \hat{\mathbb{C}}$ such that Gz is not dense in $\hat{\mathbb{C}}$.*

Proof. Note that if $\Omega = \emptyset$ then $\Lambda = \hat{\mathbb{C}}$. In this case, for all $z \in \hat{\mathbb{C}} = \Lambda$ the orbit Gz is dense in $\hat{\mathbb{C}} = \Lambda$. Conversely, if there is some $z \in \hat{\mathbb{C}}$ such that Gz is not dense in Λ then by the proposition we must have $z \notin \Lambda$, so $\Lambda \neq \hat{\mathbb{C}}$ and $\Omega \neq \emptyset$.

3.3.20 Example. Let $\zeta = \exp(2\pi i/p)$ and $\xi = \exp(2\pi i/q)$, where $p, q \in \mathbb{Z}$; for convenience, we take $p, q > 2$ (so $\sin(2\pi/p)$ and $\sin(2\pi/q)$ are both positive). Define

$$X = \begin{bmatrix} \zeta & 1 \\ 0 & \zeta^{-1} \end{bmatrix}, \qquad Y = \begin{bmatrix} \xi & 0 \\ \rho & \xi^{-1} \end{bmatrix}$$

where $\rho \in \mathbb{C}$ is non-zero. Let $G := \langle X, Y \rangle$.

The first claim is that G is non-elementary. Observe that

$$\operatorname{tr}[X,Y] = \operatorname{tr} XYX^{-1}Y^{-1} = 2 - 4\sin\frac{2\pi}{p}\sin\frac{2\pi}{q}\rho - \rho^2;$$

$[X, Y]$ is loxodromic whenever $\operatorname{tr}[X, Y] \notin [-2, 2]$. Consider the inequality

$$\left| 2 - 4\sin\frac{2\pi}{p}\sin\frac{2\pi}{q}\rho - \rho^2 \right| > 2.$$

Let α and β be the two roots of $\operatorname{tr}[X, Y]$ (as a polynomial in ρ). Then the inequality becomes

$$|\rho - \alpha||\rho - \beta| > 2$$

and so $[X, Y]$ is loxodromic whenever ρ lies on the exterior (that is, the component of the complement containing ∞) of the so-called **Cassini oval** A with locii α and β and radius $\sqrt{2}$ (Fig. 3.6).

Observe also that $\operatorname{tr} XY = 2\cos\left(\frac{2\pi}{p} + \frac{2\pi}{q}\right) + \rho$, so XY is loxodromic whenever ρ lies outside the circle C of radius 2 centred at $-2\cos\left(\frac{2\pi}{p} + \frac{2\pi}{q}\right)$.

It is easy to check that $[X, Y]$ and XY have distinct fixed points. In particular, whenever ρ lies in the common exterior of the circle C and the Cassini oval A, the loxodromic elements $[X, Y]$ and XY have between them four distinct fixed points, so $|\Lambda(G)| \geq 4$, and in particular G is non-elementary.

We consider the case $p = 3, q = 4$; the region of interest is the common exterior of the curves in Fig. 3.7.

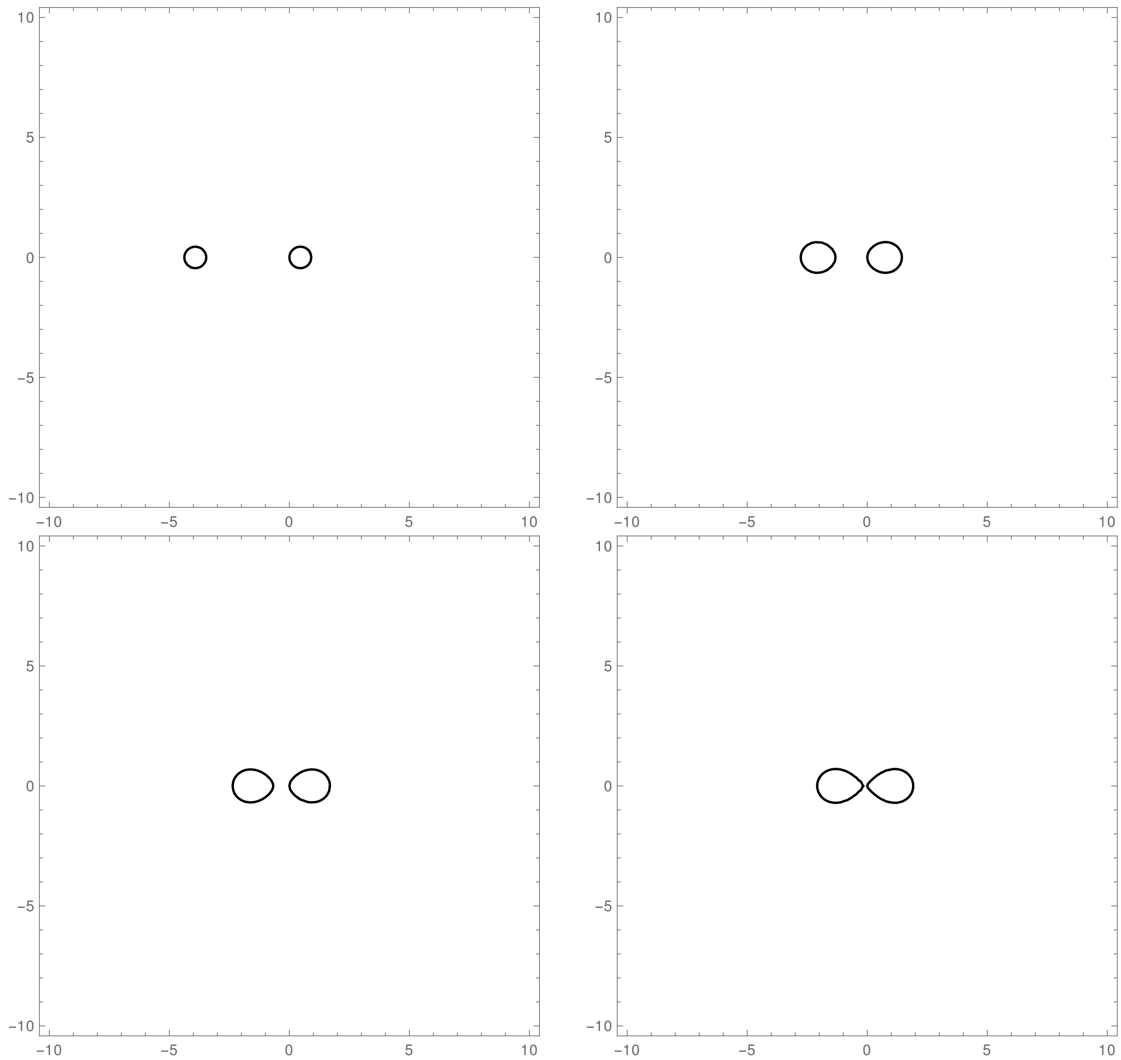


Figure 3.6: The Cassini ovals of Example 3.3.20 for $p = 3$ and $q = 4, 16, 32, 128$.

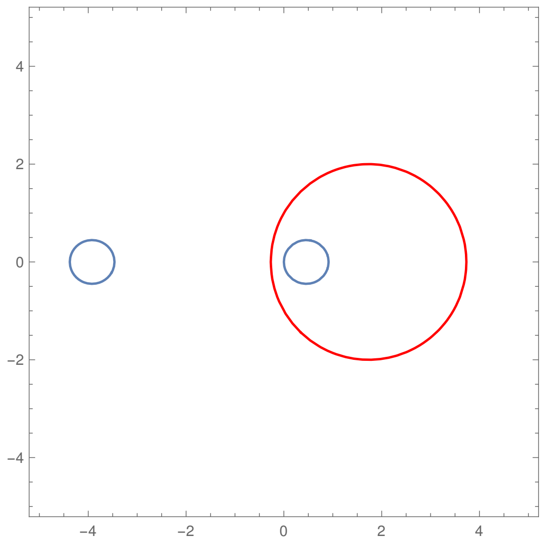


Figure 3.7: A region for which the group G is non-elementary (Example 3.3.20).

3.3.21 Theorem. *Let G be non-elementary. The set Λ may alternatively be characterised as:*

1. *The set $\overline{\Lambda_1}$, where Λ_1 is the set of loxodromic points of G;*
2. *The smallest nontrivial closed G-invariant subspace of $\hat{\mathbb{C}}$.*

Proof. 1. Let x be a loxodromic point of G, say a fixed point of $f \in G$ loxodromic. Then for almost every $z \in \hat{\mathbb{C}}$, we have $f^n z \to x$ so x is a limit point of G, i.e. $Gx \subseteq \Lambda$. By Proposition 3.3.18, Gx is dense in Λ; and since the latter is closed, $\overline{Gx} = \Lambda$. (Hence we have shown something stronger, namely Λ is the closure of the orbit of a single loxodromic point.)

2. By Theorem 3.3.4, Λ is a closed G-invariant subspace; by Proposition 3.3.16, $\Lambda \neq \emptyset$. Let E be an arbitrary such subspace. Since G is non-elementary, every orbit is infinite, and thus E is infinite. Let $x \in \hat{\mathbb{C}}$ be a loxodromic point, say fixed by $g \in G$ loxodromic. There is some $e \in E$ not fixed by g, and the set $\{g^n(e) : n \in \mathbb{Z}\}$ accumulates at x. As E is closed, $x \in E$. Thus $\Lambda_1 \subseteq E$ and again since E is closed we have $\overline{\Lambda_1} \subseteq E$.

3.3.22 Corollary. *For general G, Λ is the closure of the set Λ_0 of fixed points of non-elliptic elements.*

Proof. It is easy to see that $\Lambda_0 \subseteq \Lambda$ and hence $\overline{\Lambda_0} \subseteq \Lambda$ since the latter is closed. Let $x \in \Lambda_0$, fixed by $h \in G$ nonelliptic. Then ghg^{-1} is nonelliptic and fixes ga for all $g \in G$. In particular, Λ_0 is G-invariant. If G is non-elementary, we have $\Lambda \subseteq \overline{\Lambda_0}$ by the above theorem. On the other hand, suppose G is elementary. By Theorem 3.3.12, we have two cases:

1. $|\Lambda|$ has one element; in this case, G is of parabolic type, and by the proof of Theorem 3.3.12 this limit point is exactly the unique shared fixed point of every non-elliptic element of G; hence $|\Lambda_0| = 1$ and by counting we see $\Lambda = \overline{\Lambda_0}$ (the closure of a finite set is itself).

2. $|\Lambda|$ has two elements; in this case, G is of loxodromic type, and the limit points are exactly the two shared fixed points of the elements of G; again we see that $|\Lambda_0| = 2$, so $\Lambda = \overline{\Lambda_0}$.

Remark. Note that Beardon *defines* the limit set to be $\overline{\Lambda_1}$; so our limit set only agrees with his when the group of interest is non-elementary.

3.3.23 Corollary. *Every element of the set $\Omega \setminus {}^\circ\Omega$ is an elliptic point.*

Proof. Let $x \in \hat{\mathbb{C}}$. Suppose x is not a fixed point of any element of G; we wish to show that either $x \in {}^\circ\Omega$ or $x \in \Lambda$. Suppose $x \notin {}^\circ\Omega$; then for every neighbourhood U of x, there exists some nontrivial g such that $gU \cap U \neq \emptyset$; taking the neighbourhoods $B(x, 1/m)$ for $m \in \mathbb{N}$ we find a sequence (g_m) of elements of G and some sequence $z_m \in U$ such that $g_m z_m \to x$ and $z_m \to x$. Suppose that the g_m have an eventually constant subsequence, say equal to some element g. Then $gU \cap U \neq \emptyset$ for arbitrarily small neighbourhoods U; this implies that $gU \cap U = \{x\}$, so x is a fixed point of g — this is a contradiction. Hence we may assume that almost all the g_m are distinct. By the proof of Lemma 3.3.2 we can find a subsequence (g_n) and limit points w, y such that $g_n \to w$ uniformly on compact subsets of $\hat{\mathbb{C}} \setminus \{y\}$. If $x = y$, then $x \in \Lambda$. If $x \neq y$, then the points z_n do not accumulate at y and so the sequence $g_n z_n$ is away from y, thus $g_n z_n \to w$ and so $x = w \in \Lambda$.

In particular, we see that if x lies in $\Omega \setminus {}^\circ\Omega = \hat{\mathbb{C}} \setminus ({}^{\circ\Omega} \cup \Lambda)$ then x is a fixed point of some element. By Corollary 3.3.22, points outside Λ which are fixed points must be fixed points of elliptic elements; this completes the proof. ⅅ

The converse to Corollary 3.3.23 is not true:

3.3.24 Example. Let G be the group generated by two elements, represented by

$$X = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad Y = \begin{bmatrix} \exp(i\pi/3) & 0 \\ 0 & \exp(-i\pi/3) \end{bmatrix}.$$

Observe that G is discrete. Note also that ∞ is a fixed point of the parabolic element X, so $\infty \in \Lambda(G)$. On the other hand, $0 \in \Omega(G)$ since there are only finitely many elements of G not moving the unit disc off itself.

In particular, Y is an elliptic element with one fixed point in $\Lambda(G)$ and one fixed point in $\Omega(G)$.

3.3.25 Proposition. *$\Omega - {}^\circ\Omega$ is a discrete subset of Ω.*

Proof. Let (z_n) be a sequence of points of $\Omega - {}^\circ\Omega$. Then for each n there is a nontrivial element $g_n \in \operatorname{Stab}_G z_n$ (by Corollary 3.3.23). Each g_n lies in at most two of the $\operatorname{Stab}_G z_n$ (since transformations have at most two fixed points), so we may choose a subsequence (g_m) of distinct elements; further, by Lemma 3.3.2 choose a subsequence such that z_m converges, say to w, and such that $g_m \to x$ uniformly on compact subsets of $\hat{\mathbb{C}} \setminus \{y\}$. If $w = y$, then $w \in \Lambda$. If $w \neq y$, then the points z_m do not accumulate at y and so the sequence $g_m z_m$ is away from y, thus $g_m z_m \to x$ and so $w = x \in \Lambda$. ⅅ

3.3.26 Corollary. *Ω is non-empty iff ${}^\circ\Omega$ is non-empty.*

Proof. Since ${}^\circ\Omega \subseteq \Omega$ it suffices to show that if Ω is non-empty then ${}^\circ\Omega \neq \emptyset$; this is easy: since Ω is open in $\hat{\mathbb{C}}$, it is not discrete, so removing discrete points will not empty it. ⅅ

There is an alternative characterisation of limit points which comes from the action of G on H^3.

3.3.27 Theorem. *The set Λ is the set of all $z \in \partial H^3$ such that z is an accumulation point of Gq for some $q \in \overline{H^3}$. (In particular, the accumulation points of Gq must lie on the boundary ∂H^3.)*

Proof. This follows from the following observations: firstly, that the Poincaré extension of some $f \in \mathbb{M}$ (which is an element of GM(3)) has isometric circle the orthogonal extension of the isometric circle of f into H^3; secondly, that closures of balls in H^3 centred at a point of ∂H^3 are compact in

$\overline{H^3}$ (so the proof of Lemma 3.3.2 goes through in the relevant case, giving us $g_n \in \mathrm{GM}(3)$ which are Poincaré extensions); and hence the proof of part (1) of Theorem 3.3.3 goes through for all $q \in \overline{H^3}$, not just $z \in \partial H^3$. ♘⃗

3.3.28 Corollary. *If $q \in H^3$ and $\overline{Gq}$ denotes the closure of the orbit Gq in $\overline{H^3}$, then $\Lambda = \overline{Gq} \cap \partial H^3$.*

Proof. Observe that $\overline{Gq} \cap \partial H^3$ is precisely the set of limit points of Gq in $\overline{H^3}$, when $q \in H^3$ (not on the boundary!). ♘⃗

The reader may now use this corollary to obtain shorter/more conceptual proofs of the analysis of the limit sets of the elementary groups, following [38].

3.4 Jorgensen's inequality

3.4.1 Proposition (Shimizu-Leutbecher lemma). *Let $G \leq \mathbb{M}$ be Kleinian, containing the transformation f given by $z \mapsto z+1$. If $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ represents an element $g \in G$ distinct from f, then either $c = 0$ or $|c| > 1$.*

Proof. Assume that $|c| < 1$. Inductively define a sequence (B_m) by

$$B_0 = B$$
$$B_{m+1} = B_m A B_m^{-1},$$

and let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ represent f.

By direct computation,

$$B_m = \begin{bmatrix} a_m & b_m \\ c_m & d_m \end{bmatrix} \implies B_{m+1} = \begin{bmatrix} 1 - a_m c_m & a_m^2 \\ -c_m^2 & 1 + a_m c_m \end{bmatrix};$$

observe that $|c_m| = |c|^{2^m}$, so $|c_m| \to 0$. Since $|c_m| < 1$ for all m, we have

$$|a_{m+1}| = |1 - a_n c_n| \leq 1 + |a_n||c_m| < 1 + |a_m|$$

and so by induction $a_m < m + |a|$. Thus $a_m c_m \to 0$, so $a_{m+1} = 1 - a_m c_m \to 1$. In particular,

$$B_{m+1} = \begin{bmatrix} 1 - a_m c_m & a_m^2 \\ -c_m^2 & 1 + a_m c_m \end{bmatrix} \to \begin{bmatrix} 1 - 0 & 1 \\ 0 & 1 + 0 \end{bmatrix} = A.$$

This contradicts discreteness unless $B_m = A$ for all sufficiently large m. In this case, $c_m = 0$, so since $|c_m| = |c|^{2^m}$ we have $|c| = 0$. ♘⃗

3.4.2 Lemma. *An element $f \in \mathbb{M}$ is of order 2 iff $\operatorname{tr}^2 f = 0$.*

Proof. Suppose $\operatorname{tr}^2 f = 0$. Pick a matrix representative $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2, \mathbb{C})$ for f, then $a = -d$ and $1 = ad - bc = -a^2 - bc$; then

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^2 = \begin{bmatrix} a^2 + bc & b(a+d) \\ c(a+d) & bc + d^2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

so A^2 represents the identity map, and A is not the identity or its negation since $a = \pm 1 \implies d = \mp 1$.

Conversely, suppose $f \in \mathbb{M}$ is of order 2. Then f is elliptic, so we may pick a matrix for A of the form

$$\begin{cases} \frac{1}{x-y}\begin{bmatrix} xk^{-1} - yk & xy(k - k^{-1}) \\ k^{-1} - k & xk - yk^{-1} \end{bmatrix} & x, y \neq \infty \\ \begin{bmatrix} k^{-1} & y(k - k^{-1}) \\ 0 & k \end{bmatrix} & x = \infty \\ \begin{bmatrix} k & x(k^{-1} - k) \\ 0 & k^{-1} \end{bmatrix} & y = \infty \end{cases}$$

where $k \neq \pm 1$, and $k^4 = 1$; i.e. $k = \pm i$. In any case, $\operatorname{tr} A = 0$.

3.4.3 Example. Define the group G_ρ, for $\rho \in \mathbb{C}$, to be the group generated by the two parabolic elements

$$X = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad Y_\rho = \begin{bmatrix} 1 & 0 \\ \rho & 1 \end{bmatrix}.$$

By Proposition 3.4.1, G_ρ cannot be discrete unless $|\rho| > 1$. (This gives a bound on the **Riley slice**, see [29, §2.2].)

3.4.4 Theorem (Jørgensen's inequality)**.** *Let f, g generate a non-elementary Kleinian group. Then*

$$\left|\operatorname{tr}^2 f - 4\right| + |\operatorname{tr}[f, g] - 2| \geq 1.$$

Remark. The theorem stated as [34, theorem C.7] is a special case of the above: if f is loxodromic and g does not keep invariant the fixed point set of f then $\Lambda(\langle f, g\rangle) > 2$ and so the group $\langle f, g\rangle$ is non-elementary.

Proof. If f is of order 2, then by Lemma 3.4.2

$$\left|\operatorname{tr}^2 f - 4\right| + |\operatorname{tr}[f, g] - 2| = 4 + |\operatorname{tr}[f, g] - 2| \geq 4 \geq 1.$$

Hence we may assume f is not of order 2.

Suppose f is parabolic. By conjugation, we may assume f is the transformation $z \mapsto z + 1$; pick representatives

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

for f and g respectively. Observe that

$$\left|\operatorname{tr}^2 f - 4\right| + |\operatorname{tr}[f, g] - 2| = |\operatorname{tr}[f, g] - 2| = \left|c^2 + 2 - 2\right| = \left|c^2\right| \tag{3.4.5}$$

If $c = 0$ then g fixes ∞, and so G is elementary (e.g. by Lemma 3.3.15). In particular, we see $f \neq g$. Apply now Proposition 3.4.1; we see that $|c| > 1$, and comparison with Eq. (3.4.5) gives the result in this case.

Suppose now that f is loxodromic or elliptic. Conjugate such that the fixed points are 0 and ∞, so by the normal forms (in this case Lemma 2.3.10) we have representatives

$$A = \begin{bmatrix} k & 0 \\ 0 & k^{-1} \end{bmatrix}, \qquad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

where $k \in \mathbb{C}^*$.

If $bc = 0$, then g fixes either 0 or ∞, so $\langle f, g\rangle$ has a fixed point and is therefore elementary. Assume then that $bc \neq 0$. Observe that

$$\begin{aligned} \left|\mathrm{tr}^2 f - 4\right| + \left|\mathrm{tr}[f,g] - 2\right| &= \left|k^2 + k^{-2} - 2\right| + \left|2ad - bc(k^2 + k^{-2}) - 2\right| \\ &= \left|k - k^{-1}\right|^2 + \left|2(1 + bc) - bc(k^2 + k^{-2}) - 2\right| \\ &= \left|k - k^{-1}\right|^2 + |bc|\left|2 - (k^2 + k^{-2})\right| \\ &= (1 + |bc|)\left|k - k^{-1}\right|^2. \end{aligned} \tag{3.4.6}$$

Let μ denote this quantity. Suppose first of all that $\mu < 1$.

As in the proof of Proposition 3.4.1, inductively define a sequence (B_m) by

$$\begin{aligned} B_0 &= B \\ B_{m+1} &= B_m A B_m^{-1}, \end{aligned}$$

and now observe the relations

$$B_m = \begin{bmatrix} a_m & b_m \\ c_m & d_m \end{bmatrix} \implies B_{m+1} = \begin{bmatrix} a_m d_m k - b_m c_m k^{-1} & a_m b_m (k^{-1} - k) \\ c_m d_m (k - k^{-1}) & a_m d_m k^{-1} - b_m c_m k \end{bmatrix}; \tag{3.4.7}$$

one now notices (somehow) that

$$\begin{aligned} b_{m+1} c_{m+1} &= (a_m b_m (k^{-1} - k))(c_m d_m (k - k^{-1})) \\ &= a_m b_m c_m d_m (k^{-1} - k)^2 \\ &= (1 + b_m c_m) b_m c_m (k^{-1} - k)^2; \end{aligned}$$

we now prove by induction that $|b_m c_m| \leq \mu^m |bc|$. Indeed the base case comes from the assumption $\mu < 1$, and the inductive step is

$$\begin{aligned} |b_{m+1} c_{m+1}| &= |1 + b_m c_m||b_m c_m|\left|k^{-1} - k\right|^2 \\ &\leq |b_m c_m|\left|k^{-1} - k\right|^2 + |b_m c_m|^2\left|k^{-1} - k\right|^2 \\ &\leq \mu^m |bc|\left|k^{-1} - k\right|^2 + \mu^{2m}|bc|^2\left|k^{-1} - k\right|^2 \\ &\leq \mu^m |bc|\left|k^{-1} - k\right|^2 + \mu^m |bc|^2\left|k^{-1} - k\right|^2 \\ &\leq (1 + |bc|)\mu^m |bc|\left|k^{-1} - k\right|^2 \\ &\leq \mu^{m+1}|bc|. \end{aligned}$$

In particular, $|b_m c_m| \to 0$ and so $a_m d_m = 1 + b_m c_m \to 1$: hence

$$B_{m+1} = \begin{bmatrix} a_m d_m k - b_m c_m k^{-1} & a_m b_m (k^{-1} - k) \\ c_m d_m (k - k^{-1}) & a_m d_m k^{-1} - b_m c_m k \end{bmatrix} \to \begin{bmatrix} k & a_m b_m (k^{-1} - k) \\ c_m d_m (k - k^{-1}) & k^{-1} \end{bmatrix}$$

so $a_m \to k$, $d_m \to k^{-1}$.

Thus,

$$\left|\frac{b_{m+1}}{b_m}\right|^2 = \left|a_m (k^{-1} - k)\right|^2 \to \xi := |k|^2\left|k^{-1} - k\right|^2 \leq |k|^2(1 + |bc|)\left|k - k^{-1}\right|^2 = \mu|k|^2;$$

since $\mu < 1$, this implies that

$$\left|\frac{b_{m+1}}{b_m}\right| \to \xi \le \mu^{1/2}|k| < \frac{1+\mu^{1/2}}{2}|k|$$

so for large enough m,

$$\left|\frac{b_{m+1}}{b_m}\right| < \frac{1+\mu^{1/2}}{2}|k| \iff \left|\frac{b_{m+1}}{k^{m+1}}\right| < \frac{1+\mu^{1/2}}{2}\left|\frac{b_m}{k^m}\right| < \left|\frac{b_m}{k^m}\right|;$$

and so $|b^m k^{-m}| \to 0$.

Using similar reasoning, $|c^m k^m| \to 0$ and so

$$A^{-m}B_{2m}A^m = \begin{bmatrix} a_{2m} & b_{2m}k^{-2m} \\ c_{2n}k^{2m} & d_{2m} \end{bmatrix} \to A.$$

Since $\langle f, g\rangle$ is discrete, we must have $A^{-m}B_{2m}A^m = A$ for large enough m. This implies, by comparison between A and the expression for B_m in Eq. (3.4.7), that $a_m b_m(k^{-1} - k) = c_m d_m(k - k^{-1}) = 0$; i.e. one of a_m or b_m is zero, and one of c_m or d_m is zero. But note that $b_m c_m \neq 0$ for all m by the inductive formula above; hence $a_m = 0$ and $d_m = 0$. Thus f fixes 0 and ∞, and g swaps 0 and ∞, so the set $\{0, \infty\}$ is a finite orbit of $\langle f, g\rangle$, contradicting that this group is non-elementary. Thus $\mu \ge 1$, and the proof is completed by recalling Eq. (3.4.6).

3.4.8 Example. We show that the bound in Jørgensen's inequality is attained. Recall that the modular group $\mathrm{SL}(2, \mathbb{Z})$ is a Kleinian group generated by f and g defined by $z \mapsto z + 1$ and $z \mapsto -1/z$ respectively; then $\left|\mathrm{tr}^2 f - 4\right| + |\mathrm{tr}[f, g] - 2| = |4 - 4| + |3 - 2| = 1$.

3.5 Some examples

3.5.1 A Fuchsian group

Let $\zeta = \exp(2\pi i/p)$ and $\xi = \exp(2\pi i/q)$, where $p, q \in \mathbb{Z}$. Define

$$X = \begin{bmatrix} \zeta & 1 \\ 0 & \zeta^{-1} \end{bmatrix}, \qquad Y = \begin{bmatrix} \xi & 0 \\ \rho & \xi^{-1} \end{bmatrix}.$$

Let $G := \langle X, Y\rangle$.

We begin by computing the quantities which appear in Jørgensen's inequality (Theorem 3.4.4).

$$\mathrm{tr}^2 X = 4\cos^2\frac{2\pi}{p}$$

$$\mathrm{tr}[X, Y] = \mathrm{tr}\, XYX^{-1}Y^{-1} = 2 - 4\sin\frac{2\pi}{p}\sin\frac{2\pi}{q}\rho - \rho^2$$

In particular,

$$\left|\mathrm{tr}^2 X - 4\right| + |\mathrm{tr}[X, Y] - 2| = 4\sin^2\frac{2\pi}{p} + |\rho|\left|\rho + 4\sin\frac{2\pi}{p}\sin\frac{2\pi}{q}\right|$$

Chapter 4

Riemann surfaces

The aim of this chapter is the study of the quotient space Ω/G. A brief outline of the chapter follows:

1. In Section 4.1 we define the basic terminology and check that Ω/G is a Riemann surface;
2. In Section 4.2 we define the notion of a fundamental domain to simplify the visualisation of the Riemann surface of interest;
3. In Section 4.3 we apply this theory to find some bounds on the parabolic and elliptic Riley slices;
4. In Section 4.4 we give a method of constructing fundamental domains for certain classes of Kleinian groups;
5. In Section 7.2 we study punctures of Ω/G.

Throughout, G is a Kleinian group of the second type. If we want to talk about an abstract group, we shall use the symbol Γ.

4.1 Definitions

We recall for convenience the following definition (see [20] or [16]).

4.1.1 Definition. A **Riemann surface** is a Hausdorff space X equipped with an open cover $(U_\alpha)_{\alpha\in A}$ and, for each $\alpha \in A$, an injective open map $\phi_\alpha : U_\alpha \to \mathbb{C}$ (called a **coordinate chart**), such that for all $\alpha, \beta \in A$ the **transition map**

$$\phi_\alpha\phi_\beta^{-1} : \phi_\beta(U_\alpha \cap U_\beta) \to \phi_\alpha(U_\alpha \cap U_\beta)$$

is holomorphic.

A map $f : X \to Y$ (X and Y Riemann surfaces) is **holomorphic** if for each pair of charts ϕ for X and ψ for Y, the composition $\psi f \phi^{-1}$ is holomorphic. We say that such an f is a **biholomorphism**, a **conformal equivalence**, or simply an **equivalence** if f is invertible and f^{-1} is holomorphic; we write $X \simeq Y$ in this case.

A Riemann surface is, in particular, a 2-manifold; the second-countability of a Riemann surface is nontrivial and is known as **Rado's theorem**; see the remark on page 140 of [16]. A Riemann surface which is a 2-manifold that is *compact and without boundary* is called **closed**.

Let X and Y be Riemann surfaces, with $p : X \to Y$ a non-constant holomorphic map. Then p is open (being open is a local property, and holomorphic maps between subsets of $\mathbb{C}$ are open), and $p^{-1}(y)$ is discrete for any $y \in Y$ (again, this follows from the fact that locally a holomorphic map is constant if the inverse image of any point $y \in Y$ has a limit point).

A point $x \in X$ is a **branch point** or **ramification point** of p if there is no neighbourhood of x on which p is injective.

We will be interested in Riemann surfaces which have a 'hyperbolic structure'. We say that a 2-manifold M is **hyperbolic** if there is an atlas $(\phi_\alpha : U_\alpha \to B^2)_{\alpha \in A}$ on M (where B^2 is viewed as a subset of $\mathbb{R}^2$) with the following properties:

1. Each U_α is connected;
2. If $U_\alpha \cap U_\beta \neq \emptyset$, then for every $x \in U_\alpha \cap U_\beta$, there is a neighbourhood $U \ni x$ and an element $g \in \mathbb{M}$ such that $(\phi_\beta \phi_\alpha^{-1}) \restriction_U = g$.

If X is a Riemann surface such that the Riemann surface structure is also a hyperbolic structure under the identification of B^2 with the ball model of hyperbolic space in $\hat{\mathbb{C}}$, then we say that X is a **hyperbolic Riemann surface**.

Let a group Γ act on a set X, and let $H \leq \Gamma$. We say that X is **precisely invariant** under H if the following conditions hold:

1. $H = \Gamma_X$, and
2. $gX \cap X = \emptyset$ for all $g \in \Gamma \setminus H$.

If X is precisely invariant under the identity, then we say X is a G-**packing**.

4.1.2 Lemma. *A point x lies in $\Omega(G)$ iff the following conditions hold:*

1. $\mathrm{Stab}_G\, x$ *is finite; and*
2. *there is a neighbourhood $U \ni x$ which is precisely invariant under* $\mathrm{Stab}_G\, x$.

Proof. If G acts discontinuously about $x \in \hat{\mathbb{C}}$, it is clear that $\mathrm{Stab}_G\, x$ is finite. Pick a neighbourhood U of x such that $gU \cap U$ is nonempty for only finitely many $g \in G$, say $g_1, \dots, g_k$. For each g_i not a stabiliser of x, there exists a sufficiently small disc $B_i = B(x, \epsilon_i)$ such that $d(g_i B_i, x) > 0$, and for all $g_i \in \mathrm{Stab}_G\, x$ set $B_i = U$; now replace U with $U \cap \bigcap_i B_i$; then $gU \cap U$ is nonempty iff $g \in \mathrm{Stab}_G\, x$. It follows that $U^* = \bigcap_{g \in \mathrm{Stab}_G x} g(U)$ is a neighbourhood of x precisely invariant under $\mathrm{Stab}_G\, x$.

The converse is trivial. ∎

4.1.3 Theorem. *The quotient space $\mathcal{R}_G \coloneqq \Omega/G$ is a hyperbolic Riemann surface.*

Proof. Note that $\mathcal{R}_G$ is Hausdorff by the theory of Section 3.1, and that the projection π is open by standard considerations [33, chapters 11 and 12].

We now define the complex structure on $\mathcal{R}_G$. In fact, we will show that there is an atlas with charts formed by the projections of nice discs about each $x \in \Omega$ onto $\mathcal{R}_G$ with the G-transition property.

Let D be the unit disc in $\mathbb{C}$.

Suppose first that $z \in {}^\circ\Omega$; we may pick a neighbourhood U_z of z such that $gU_z \cap U_z \neq \emptyset$ only for $g = 1$. Then the restriction $\pi_z \coloneqq \pi \restriction_{U_z} : U_z \to \Omega/G$ is injective and so there exists $\pi_z^{-1} : \pi(U_z) \to U_z$ a left inverse for π_z on some neighbourhood of $\pi(z)$. Let $\sigma : U_z \to D$ be a surjective Möbius transformation sending $z \mapsto 0$; then $\varphi_z \coloneqq q\sigma\pi_z^{-1} : \pi(U_z) \to D$ is a homeomorphism (where q is the identity map).

On the other hand, suppose $z \in \Omega \setminus {}^\circ\Omega$. Let U_z be a neighbourhood of z precisely invariant under G_z (which exists by Lemma 4.1.2); and let $J = G_z$. Note that there is a natural homeomorphism

$U_z/J \simeq U_z/G$ by the properties of precise invariance. Observe that J is finite and so each nontrivial element of J is elliptic. We note that, by Corollary 3.3.23, such a nontrivial element exists. Suppose $f, g \in J$; then $[f, g] \in J$ is elliptic and so by Lemma 2.4.2 we have that $\operatorname{Fix} f = \operatorname{Fix} g$. Let σ be a Möbius transformation sending $z \to 0$ and U_z onto the unit disc D. Since J is finite and consists of elliptic elements, it is cyclic and so is generated by some elliptic $g \in J$ satisfying $\sigma g \sigma^{-1}(w) = \exp(2\pi i/n)w$ for all $w \in D$. Let $q : D \to D$ be the map $z \mapsto z^n$. For all $w \in U_z$ and all $k \in \mathbb{Z}$ we have

$$q\sigma g^k w = \left(\sigma g^k \sigma^{-1} \sigma w\right)^n = \left(\exp(k 2\pi i/n)\sigma w\right)^n = (\sigma w)^n.$$

Observe that this shows that the function $q\sigma\pi_z$ (where $\pi_z := \pi\restriction_{U_z}$) is well-defined, regardless of the branch of the inverse that is chosen: indeed, $\pi(w)$ is mapped by a branch of π_z^{-1} to some point $g^k(w)$ and the image of each of these under $q\sigma$ is $(\sigma w)^n$. In particular, $\varphi_z := q\sigma\pi_z : \Omega/G \supseteq \pi(U_z) \to D$ is a homeomorphism.

Let $\Sigma = \{(\varphi_z, \pi(U_z)) : z \in \Omega\}$; the claim is that this is a chart for $\mathcal{R}_G$.

We begin by studying the maps $\pi_z^{-1}\pi_w$ for $w \neq z$. Let $\zeta_w \in U_w$, $\zeta_z \in U_z$. If $\pi(\zeta_w) = \pi(\zeta_z)$ then for some $g \in G$, we have $g\zeta_w = \zeta_z$. If ζ_w and ζ_z are not elliptic fixed points, then π_z^{-1} exists (is univaluate) in some neighbourhood of ζ_z and takes values in $\pi_z(U_z)$. The two maps $\pi_z g$ and π_w with π on some neighbourhood of ζ_w; thus on that neighbourhood, $g = \pi_z^{-1}\pi_w$. Thus the maps $\pi_z^{-1}\pi_w$ are elements of G away from elliptic fixed points of G.

We may write $\varphi_z = q_z\sigma_z\pi_z^{-1}$ and $\varphi_w = q_w\sigma_w\pi_w^{-1}$; suppose $\zeta \in \pi(U_z) \cap \pi(U_w)$. If ζ is not an elliptic fixed point, then we may pick an arbitrary branch of q_w^{-1} about ζ (if necessary) and then compute

$$\varphi_z\varphi_w^{-1} = q_z\sigma_z\pi_z^{-1}\pi_w\sigma_w^{-1}q_w^{-1};$$

If ζ is not an elliptic fixed point, by the previous paragraph we have

$$\varphi_z\varphi_w^{-1} = q_z\sigma_z\pi_z^{-1}\pi_w\sigma_w^{-1}q_w^{-1} = q_z\sigma_z g\sigma_w^{-1}q_w^{-1}$$

which is clearly a biholomorphism; on the other hand, if ζ is an elliptic fixed point then we may pick a punctured neighbourhood of ζ on which $\varphi_z\varphi_w^{-1}$ is biholomorphic by the previous remark (this is because elliptic fixed points are discrete) and then by standard considerations (Hartog's theorem) it may be extended to a biholomorphic function defined at ζ.

Thus we have shown that $\mathcal{R}_G$ is a Riemann surface with charts being homeomorphisms onto D. It is a consequence of Schwartz' lemma ([2, section 3.4, exercise 5]) that such maps are in fact Möbius transformations which fix D, and thus act as hyperbolic isometries in D.

4.1.4 Lemma. *Let X be a compact 2-manifold. Every connected component W of the boundary has a neighbourhood U with interior $U \setminus W$ homeomorphic to an annulus such that the homeomorphism extends continuously to $U \cap W$ and maps $\bar{U}$ continuously onto a closed annulus; in particular, each boundary component of a Riemann surface has a neighbourhood conformally equivalent to a punctured disc or an annulus.*

If W is a boundary component of a Riemann surface with a neighbourhood conformally equivalent to a punctured disc, we say that W **bounds a puncture** or that the surface is **punctured** at W; otherwise, it **bounds a disc** or corresponds to a **removed disc**.

The Riemann surface $\mathcal{R}_G$ admits a marking structure as follows (this is a definition, not a theorem):

- projections of elliptic elements $x \in \Omega$ are marked with the order of the generator of $\operatorname{Stab}_G x$;
- punctures are marked with ∞.

A Riemann surface X is said to be **of finite type** if X is biholomorphic to a compact Riemann surface with at most finitely many points removed.

4.1.5 Theorem (Ahlfors' finiteness theorem). *If G is a finitely generated non-elementary Kleinian group, then $\mathscr{R}_G$ has finitely many components, each hyperbolic of finite type. In particular, if* $\operatorname{area}(\mathscr{R}_G)$ *denotes the hyperbolic area of $\mathscr{R}_G$, then* $\operatorname{area}(\mathscr{R}_G)$ *is finite.*

Historical remark. The theorem was first published in [4], an error in the proof was indicated by Bers [3], and the gap was filled in [23].

When we study the structure of hyperbolic 3-manifolds, we shall provide a proof of Ahlfors' theorem following [31].

There are some useful quantitative extensions to Ahlfors' finiteness theorem. For instance:

4.1.6 Theorem (Bers, 1967). *If G is a finitely generated non-elementary Kleinian group, then*

$$\frac{1}{2\pi} \operatorname{area}(\mathscr{R}_G) \leq 2(N-1)$$

where N is the size of a minimal generating set for G.

Finally we remark that there is a converse to the construction of the Riemann surface $\mathcal{R}_G$ from a Kleinian group G.

4.1.7 Theorem (Klein-Poincaré Uniformisation Theorem). *Let R be a Riemann surface, and let $\hat{R}$ be its universal cover. Then $\hat{R}$ is conformally equivalent to $\hat{\mathbb{C}}$, $\hat{\mathbb{C}}$, or $B^2 = \{z \in \mathbb{C} : |z| < 1\}$; and $R = \hat{R}/G$, where G is a Kleinian group acting discontinuously on $\hat{R}$ (where the action is induced by restriction); further, G is unique up to conjugation in* $\operatorname{Aut}\hat{R}$.

Motivated by this, we define a **Fuchsian group** to be a Kleinian group which acts discontinuously on some disc D (i.e. a conformal image of B^2).

In fact, we have a stronger result:

4.1.8 Theorem (Bers' simultaneous uniformisation theorem, 1960). *Let S and S' be quasiconformally equivalent hyperbolic Riemann surfaces with degenerate boundary components; then there exists a quasi-Fuchsian group of the first kind such that $\Omega(G)/G = S \cup S'$.*

An excellent discussion of these results and many others together with a plethora of historical references is found in [30].

4.2 Fundamental domains

Our next goal is the study of the quotient $\mathcal{R}_G$. In order to do this, we will find a region $D \subseteq \Omega$ such that the action of G on D tiles the entirety of Ω, and such that D is the 'smallest possible' such region.

More precisely, we make the following definition.

4.2.1 Definition. A **fundamental domain** for G is an open set $D \subseteq \Omega$ with the following properties:

1. D is a G-packing;
2. For every $z \in \Omega$, there exists $g \in G$ such that $gz \in \overline{D}$;
3. The boundary of D consists of limit points of G and a countable collection of curves $(\gamma_i : [0,1] \to \hat{\mathbb{C}})_{i \in I}$ such that $\partial D = \cup_{i \in I} \gamma_i[0,1]$ and such that $\gamma_i(0,1) \subseteq \Omega$ for all i; the intersections $s_i := \gamma_i[0,1] \cap \Omega$ are called the **sides** of D;

4. If s is a side of D, then there exists a side s' and a nontrivial element $g \in G$ (called a **side-pairing transformation**) such that $s' = gs$, and these choices are made such that $g = (g')'$;
5. If (s_n) is a sequence of sides of D, then diam $s_m \to 0$ and the sides of D accumulate only at limit points;
6. Only finitely many translates of D meet any compact subset of Ω.

Some unfortunate terminology: a fundamental domain is *not* a domain (it is not necessarily simply connected).

Observe that condition (1) in the above is equivalent to the statement that "$\forall_{x,y\in D}(gx = y) \implies g = 1$" (that is, if $x \in D$ then x has no nontrivial G-translates in D). Indeed, suppose D is precisely invariant under 1; if $gx = y$ for $g \in G$ then $y \in gD \cap D$ and so $g = 1$. On the other hand, suppose $gD \cap D$ is nonempty; then an element in the intersection is a nontrivial translate in D.

Endpoints of sides of D which lie in Ω are called **vertices** for D.

4.2.2 Lemma. *Let D be a fundamental domain, and let g be a side-pairing transformation for G, pairing $s \mapsto s'$. Then $\overline{gD} \cap \overline{D} \cap \Omega$ is a union of sides of D containing s'.*

Proof. Clearly $s' \subseteq \overline{gD} \cap \overline{D} \cap \Omega$, by definition. Suppose $x \in \overline{gD} \cap \overline{D} \cap \Omega$. Observe that if x lies in the interior of D, then x lies in the interior of gD. Hence $D \cap gD$ is nontrivial, contradicting that D is a G-packing. Thus x is a boundary point of D in Ω, and hence is an element of a side. Finally note that if an element of a side s lies in $\overline{gD} \cap \overline{D} \cap \Omega$ then the entire side lies in the set, and thus the set is a union of sides. ∎

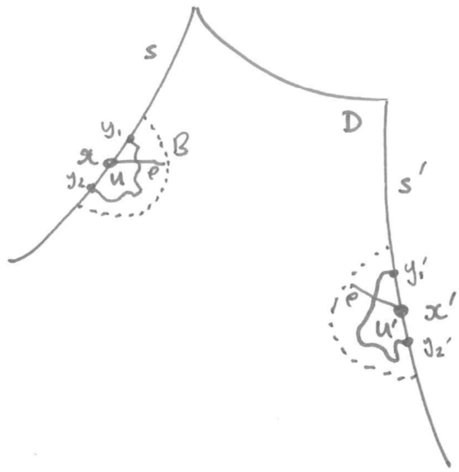
We say that D *tesselates* Ω. We write $\tilde{D}$ for $\overline{D} \cap \Omega$; the action of G on the sides of $\tilde{D}$ induces a gluing of the sides of $\tilde{D}$. In fact:-

4.2.3 Theorem. *Let $p : \tilde{D} \to \Omega/G$ be the restriction of the projection map $\Omega \to \Omega/G$. It is clear by definition of the action of G on $\tilde{D}$ that the projection induces a map $\tilde{p} : \tilde{D}/G \to \Omega/G$. Then $\tilde{p}$ is an homeomorphism.*

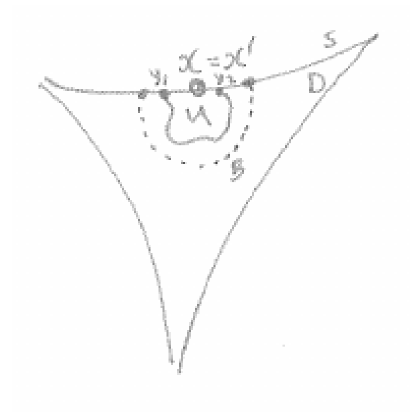
Proof. By condition (2), the map $\tilde{p}$ is surjective. By condition (1), it is injective on D. Observe that $\tilde{p}\restriction_D$ is a local homeomorphism: let $x \in D$ and let U be a neighbourhood of x in D; then $\tilde{p}(U) = p(U)$ is open since p is open on D. It remails to show that $\tilde{p}$ is a local homeomorphism on $\tilde{D}/G$ and is injective.

Case 1. Suppose first that x is an interior point of a side s of D. Pick a $g \in G$ with the property $gs = s'$ for some side s'. Let $x' = g(x)$.

> **Case I:** $x \neq x'$. Let δ (resp. δ') be the minimal distance from x (resp. x') to x' (resp. x), any vertex of D, any limit point of G, or any fixed point of g. By condition (5), both δ and δ' are nonzero. Choose $\varepsilon < \min\{\delta, \delta'\}/2$ small enough such that $B = B(x, \rho)$ is precisely invariant under $\mathrm{Stab}_G\, x$. Choose $y_1, y_2 \in B \cap s$ such that y_1 and y_2 lie on opposite sides of x in s. Let γ be a path connecting y_1 and y_2 that, except for its endpoints, lies in $B \cap D$. This path defines a closed neighbourhood U of x in $\tilde{D}$. Similarly, we may define a neighbourhood U' of x' (Fig. 4.1a).
>
> Note that each point of $D \cap U$ (resp. $D \cap U'$) is equivalent in $\tilde{D}$ only to itself. Further, every point of $s \cap U$ is equivalent only to the corresponding point of $s' \cap U$. Let $V = U \cap g^{-1}(U')$. This is a set containing x; further, it is precisely invariant under $\mathrm{Stab}_G\, x$ since B is. Since no points of $V - s(\subseteq D)$ are G-equivalent, no points of V are G-equivalent. Hence $\tilde{p}$ restricted to $U \cup U'$ is a homeomorphism, i.e. $\tilde{p}$ is a local homeomorphism about x.



(a) Case I.



(b) Case II.

Figure 4.1: Definition of U and U' in case 1 of Theorem 4.2.3.

Case II: $x = x'$. If $gx = x$, then $g = g^{-1}$ (by (4)). Let δ be the minimal distance from x to any vertex of D, to any other fixed point of g, or to any limit point of G. As above choose $\varepsilon < \delta/2$ small enough such that $B = B(x, \rho)$ is precisely invariant under $\mathrm{Stab}_G\, x$. Let $y_1 \in B \cap s$ distinct from x and set $y_2 = gy_1$. Again we obtain a closed neighbourhood U of x (Fig. 4.1b).

Observe that: the points of $U \cap D$ are G-equivalent only to themselves; a point $z \in s \cap U$ is equivalent only to gz, which also lies in U; and x is equivalent only to itself. Let $V = U \cup gU$. Then V is precisely invariant under the identity and $\tilde{p}\restriction_V$ is a homeomorphism.

Case 2. The second case is that $x = x_1$ is a vertex of D. Let s_1 be a side of D with x_1 as an endpoint. Let s_1' be the side paired with s_1, let g_1 be the side-pairing transformation relating s_1 to s_1', and let $x_2 = g_1x_1$; let s_2 be the side distinct from s_1' with x_2 as an endpoint. Inductively construct sequences (x_i), (s_i), (g_i), and (s_i'). Note that by property (6), this process ends after only finitely many steps. Let n be the smallest integer greater than 1 with $s_n = s_1$. Choose points y_i on each s_i as in case 1, and for all i set $y_i' = g_iy_i \in s_i'$. For each i, choose a path γ_m from y_{m-1}' to y_m, so that the open region U_m bounded by γ_m, s_{m-1}', and s_m lies in D. The closure of each U_m is precisely invariant under $\mathrm{Stab}_G\, x_m$, and the U_m are all disjoint. Let U be the projection in $\tilde{D}/G$ of the union of the G_m; then U is a neighbourhood of the projection of x.

Observe, $g_1^{-1}(D)$ abuts D along s_1; $g_1^{-1}g_2^{-1}(D)$ abuts $g_1^{-1}(D)$ along $g_1^{-1}(s_2)$; etc. The union of the sets U_1, $g_1^{-1}(U_2)$, $g_1^{-1}g_2^{-1}(U_3)$ need not be a neighbourhood of x: the element $h \coloneqq g_{n-1}^{-1}\cdots g_1^{-1}$ may be a nontrivial element of $\mathrm{Stab}_G\, x$ mapping s_1 onto some other arc $h(s_1)$ eminating from x. Since no two points of D are G-equivalent, either $h^m(V)$ is disjoint from V, or is

Thus the gluing of D induced by G does give $\mathcal{R}_G$.

4.2.4 Corollary. *If D is a fundamental domain for G, and if $z \in \bar{D}$ is the preimage of a marked point of $\mathcal{R}_G$, then either z is a vertex of D, or z is the fixed point on some side s of a side pairing transformation g with $g(s) = s$.*

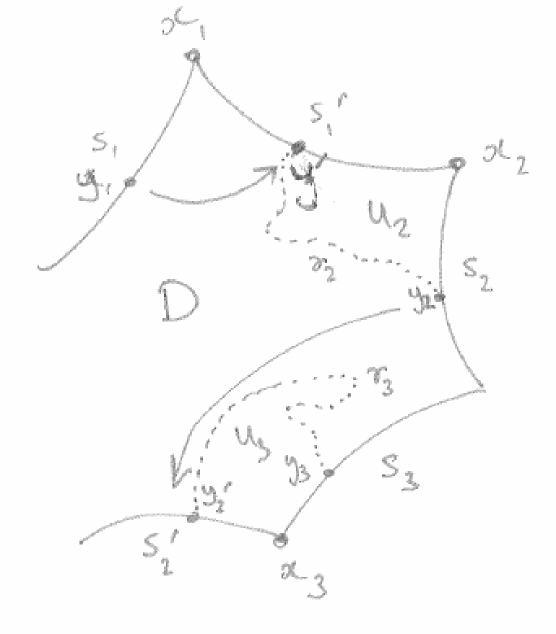


Figure 4.2: Definition of U in case 2 of Theorem 4.2.3.

The construction of fundamental domains is, in general, difficult; the difficulty is usually property (1). One useful result is the following

4.2.5 Theorem (Klein combination theorem). *Let $G_1, G_2, \ldots$ be Kleinian groups, with $G = \langle G_1, \ldots \rangle$. For each j let D_j be a G_j-packing, and suppose that $D_i \cup D_j = \hat{\mathbb{C}}$ for all $i, j \in \mathbb{N}$, $i \neq j$; suppose also that $D^* := \bigcap_{j \in \mathbb{N}} D_j \neq \emptyset$. Then G is the free product of the G_j, D^* is a G-packing, and G acts discontinuously on $\cup_{g \in G} g(D^*)$.*

Proof. Beardon, thm 5.3.15.

We have a more general theorem for Fuchsian groups which we shall prove later on (Theorem 9.3.2).

4.3 Bounds on Riley slices

We now place in context the brief discussions above, of Example 3.1.13 and Example 3.4.3. We follow essentially [29, section 2.2].

Let $\langle f, g \rangle$ be a Kleinian group, free on the two parabolic generators f and g. In order to study the group, we conjugate f to have a fixed point at ∞ and a translation length of 1, and conjugate g to have a fixed point at 0. Thus we may pick representatives for f and g of respective forms

$$X = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad Y_\rho = \begin{bmatrix} 1 & 0 \\ \rho & 1 \end{bmatrix};$$

we write G_ρ for $\langle X, Y_\rho \rangle$. We shall see later (c.f. [29, top of p.75]) that if $[X]$ and $[Y_\rho]$ are the maximal parabolic conjugacy classes of G_ρ, then $\mathcal{R}_{G_\rho}$ is a 4-times punctured sphere (with certain additional structure).

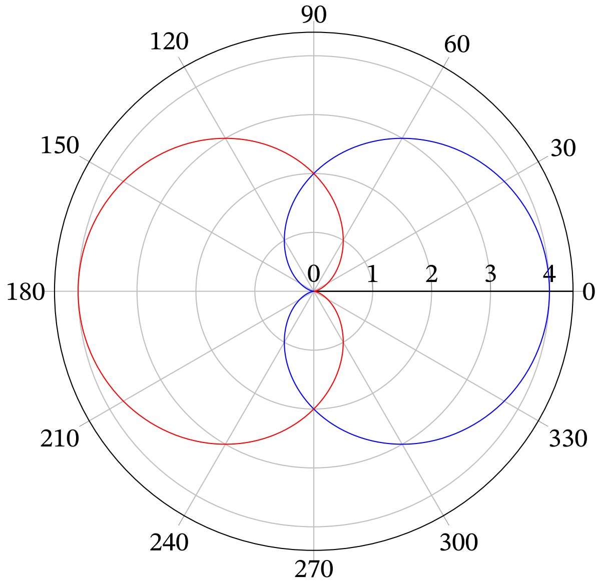


Figure 4.3: The cardioids of Lemma 4.3.1.

We define the (parabolic) **Riley slice** to be the moduli space of such groups:

$$\mathscr{R}_{\text{par.}} := \{\rho \in \mathbb{C} : \mathcal{R}_{G_\rho} \text{ is a 4-times punctured sphere}\}.$$

In Example 3.4.3, we obtained a rough bound on $\mathscr{R}_{\text{par.}}$; namely, the space lies in the exterior of the unit circle.

4.3.1 Lemma. *If ρ lies in the common exterior of the cardioids*

$$\{r\exp(i\theta) \in \mathbb{C} : r = \pm 2(1+\cos\theta)\},$$

(see Fig. 4.3) then the isometric circles of Y_ρ and Y_ρ^{-1} lie in the strip

$$S = \{z \in \mathbb{C} : -\frac{1}{2} < \operatorname{Re} z < \frac{1}{2}\}.$$

The region D inside the strip and outside the circles is a fundamental domain for G_ρ.

Remark. This lemma shows that each such G_ρ is a **classical Schottky group** of rank 2.

Proof. Suppose $\rho = r\exp(i\theta)$ lies outside both cardioids; then

$$r > |2(1+\cos\theta)|. \tag{4.3.2}$$

Let $I(Y_\rho)$ denote the isometric circle of Y_ρ, so $I(Y_\rho) = S(-\rho^{-1}, r^{-1})$; we have that $\operatorname{Re}(-\rho^{-1}) = -r^{-1}\cos\theta$. In particular, $I(Y_\rho)$ lies in the strip S iff the following inequalities hold:

$$-\frac{1}{2} < -r^{-1}\cos\theta - r^{-1} = -r^{-1}(1+\cos\theta) \iff r > 2(1+\cos\theta)$$
$$\frac{1}{2} > -r^{-1}\cos\theta + r^{-1} \iff r > 2(1-\cos\theta)$$

The first inequality follows immediately from Eq. (4.3.2); the second inequality holds since after observing that the cardioid described by $r = -2(1 + \cos\theta)$ is identitical to that described by $r = 2(1 - \cos\theta)$. Similar inequalities hold for $I(Y_\rho^{-1})$.

We now check that the region D is a fundamental domain. It is clearly open, in $\hat{\mathbb{C}}$ by construction. That it is a subset of Ω (and hence open in Ω) will follow from property (1) of Definition 4.2.1. Indeed, for $x \in D$ let U be a ball about x; if $gU \cap U \neq \emptyset$ then $gD \cap D \neq \emptyset$, and hence by the cited property we have $g = 1$. We now check properties (1) to (6) of Definition 4.2.1.

1. This follows from Theorem 4.2.5 with $G_1 = \langle X \rangle$ and $D_1 = S$, and $G_2 = \langle Y_\rho \rangle$ and $D_2 = I(Y_\rho) \cup I(Y_\rho^{-1})$.

2. For any $z \in \Omega$, there is some $n \in \mathbb{N}$ such that $X^n z \in \overline{S}$; if $X^n z$ lies in

Remark. In fact, $D = \partial H^3 \cap \bar{E}$, where E is the *Dirichlet region* for G_ρ, as defined in the chapter on 3-manifolds.

4.4 The Ford region

Suppose now that G is a Kleinian group with $\infty \in {}^{\circ\circ}\Omega$; in particular, ∞ is not a fixed point of any element. For each nontrivial $g \in G$, let D_g be the exterior of the isometric circle of g (that is, the component of the complement of the isometric circle which contains ∞). Then the **Ford region** of G is the set $D \coloneqq \operatorname{int} \bigcap_{g \in G \setminus \{1\}} \overline{D_g}$.

4.4.1 Theorem. *The Ford region is a fundamental domain for G.*

Proof. The set D is open by construction; condition (1) of Definition 4.2.1 then shows that $D \subseteq \Omega$ (in fact $D \subseteq {}^{\circ}\Omega$). We verify now the conditions of Definition 4.2.1.

1. Each nontrivial $g \in G$ sends the exterior of $I(g)$ into the interior of $I(g^{-1})$; hence gD lies in the interior of $I(g^{-1})$ and $gD \cap D = \emptyset$.

2. Let $z \in \Omega$. It follows from Corollary 3.2.5 that z lies in D_g for all but finitely many $g \in G$ (otherwise z is a limit point, namely the limit of ∞ under some sequence of elements (g_n) of G with $z \in D_{g_n}$).

Chapter 5

Palate cleanser: convexity theory in hyperbolic space

In this chapter, we gather various results about convex subsets of hyperbolic space. We essentially follow a mixture of [43, chapter 8], and the various papers collected in [17, 12].

5.1 The convex hull in general

We give some standard definitions and results which may be found in [18].

5.1.1 Definition. A subset $X \subseteq \overline{H^n}$ is **hyperbolically convex** or simply **convex** if, for any pair of points $x, y \in X$, the geodesic arc $[x, y]$ lies in X. If X is, in addition, closed, a **supporting hyperplane** for X is a hyperbolic hyperplane P such that $P \cap X \neq \emptyset$ and such that X lies entirely in one of the closed half-spaces determined by P. The half-space containing X is then labelled P^+ and is called a **supporting half-space** for X; the other half-space determined by P is labelled P^-. A **face** of X is an intersection $X \cap P$, where P is a supporting hyperplane for X.

Let $\Lambda \subseteq \overline{H^n}$ be closed; the **convex hull** of Λ, denoted h−conv Λ, is the intersection of the convex sets containing Λ.

See Fig. 5.1 for some examples.

Observe that this definition is essentially the same as that for affine space (compare for instance [19, chapter 1]). The analogy is very strong, in particular we have a retraction map with similar properties to the affine retraction onto a convex set. This analogy may be carried further, giving a general axiomatic theory of convexity; see for instance [27].

5.1.2 Theorem. *Let $X \subseteq \overline{H^n}$ be a closed convex set. Define the* ***retraction*** $r : \overline{H^3} \to X$ *by*

$$ry := \begin{cases} y & y \in X; \\ \text{the point of intersection of the largest hyperbolic sphere centred at } y \text{ with interior disjoint from } X & y \in H^3 \setminus X; \\ \text{the point of intersection of the largest horosphere based at } y \text{ with interior disjoint from } X & y \in \partial H^3. \end{cases}$$

Then r is a continuous function such that for all $x, y \in H^n$, $d(rx, ry) \leq d(x, y)$ (we say that r is ***distance-decreasing****).*

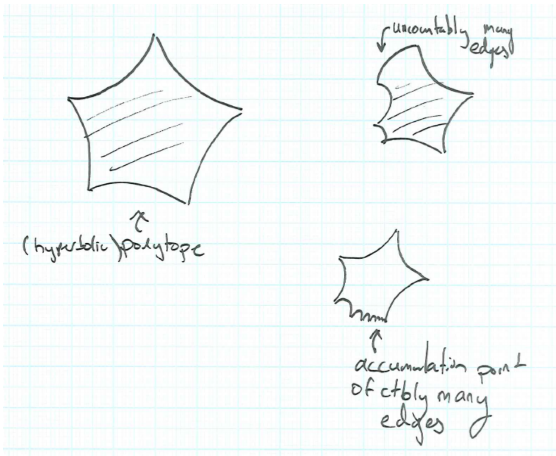


Figure 5.1: Some hyperbolic convex sets, illustrating the variety allowed.

Proof. It is obvious by considering, for instance, the ball model of $\overline{H^n}$ on which the Euclidean topology and the hyperbolic topology agree, that r is continuous. If $rx \neq ry$, then let λ be the oriented line segment from ry to rx and let L be the oriented line determined by λ. Let $\pi : H^n \to L$ be the orthogonal projection map; then $\pi x \leq rx < ry \leq \pi y$ (indeed, it cannot be that $\pi x > rx$: if $rx < \pi x \leq ry$ then $\pi x \in X$ and so πx is a point of X closer to x than rx, and if $ry < \pi x$ then ry is a point of X closer to x than rx; similarly we cannot have $\pi y < ry$) and so since π is distance-decreasing we have $d(rx, ry) \leq d(\pi x, \pi y) \leq d(x, y)$.

5.1.3 Proposition. *Given any $y \in \overline{H^n} \setminus X$, where X is a closed convex set, there exists a support plane P for X such that $y \in P^-$.*

Proof. Let $S(y)$ be the sphere (either a hyperbolic sphere or a horosphere) which defines the value of the retraction $x := ry$ as in the displayed equation of Theorem 5.1.2. Let P be the tangent plane to $S(y)$ at x. We now show that P is a supporting hyperplane; it trivially has nontrivial intersection with X and so it remains to show that X lies entirely in one of the closed half-spaces determined by P. Suppose not; then there exists some $x' \in X$ which lies on the same side of P as y. Let Q be the plane spanned by x, x', y, and let E be the hyperbolic circle $Q \cap S(y)$ (Fig. 5.2). Observe that, since P is tangent to $S(y)$, the segment $[x, x']$ (which is contained in X by convexity) passes through the interior of E, and so there is some $z \in X$ with $d(y, z) < d(y, x)$, a contradiction.

5.1.4 Corollary. *A closed $X \subseteq \overline{H^n}$ is the intersection of its supporting half-spaces.*

Proof. Suppose $y \notin X$; then $y \in P^-$ for some supporting plane by Proposition 5.1.3 and thus y does not lie in the intersection of the supporting half-spaces.

5.1.5 Corollary. *A closed $X \subseteq \overline{H^n}$ is the intersection of* countably many *of its supporting half-spaces (hence has only countably many facets).*

Proof. Let $\{y_i\}$ be a countable dense subset of $\overline{H^n} \setminus X$; then it suffices to take seperating planes for the y_i only in the proof of the previous Corollary 5.1.4. (Indeed, any element x not in X may be approximated by a sequence of the $\{y_i\}$, and the sequence of seperating planes generated by the subsequence will in the limit be a seperating plane for x.)

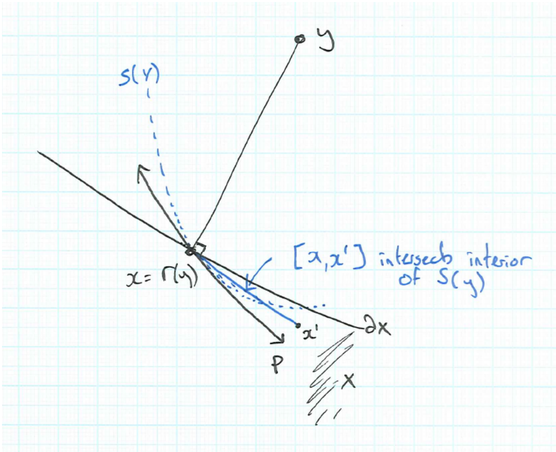


Figure 5.2: Illustration for the proof of Proposition 5.1.3.

The above Corollary 5.1.5 shows that a convex hull of a closed set will consist of countably many ‘flat’ pieces seperated by possibly uncountably many ‘pleats’ or ‘bending lines’; refer back to Fig. 5.1 for an illustration of these.

A **hyperbolic subspace** of H^n of dimension m $(0 \le m \le n)$ is the intersection (in the ball model) of a Euclidean sphere S^m or m-dimensional hyperplane orthogonal to S^{n-1} with H^n. If $X \subseteq H^n$, then the smallest hyperbolic subspace containing X is known as the **span** of X. The (topological) interior of X as a subset of its span is the **relative interior** relint X. We define the **dimension** of a convex set is the dimension of the hyperbolic subspace it spans. We say that a closed convex $X \subseteq \overline{H^n}$ is **full** if $\dim X = n$. If a face of X is 1-dimensional, we call it an **edge**. If a face has dimension $\dim X - 1$, we call it a **facet**.

We say that m points in H^n are **in general position** if they span a hyperbolic subspace of dimension m. A m-**simplex** is the hyperbolic convex hull of m points in general position.

5.1.6 Lemma. *Let* $\Delta = \text{h}-\text{conv}\{x_1, \dots, x_m\}$ *be a hyperbolic* m*-simplex spanned by the* x_i. *Then* relint $\Delta \neq \emptyset$.

Proof. Proceed by induction on m; the smallest nontrivial case is the case $m = 2$, in which case the points x_1 and x_2 are distinct points and so $[x_1, x_2]$ has non-empty relative interior. In the general case, there is a point $y \in \text{relint}\, \text{h}-\text{conv}\{x_1, \dots, x_{m-1}\}$; then $[x_m, y]$ lies in relint h$-$conv$\{x_1, \dots, x_m\}$. ⅅ

5.1.7 Theorem. *Let X be a full closed convex subset of $\overline{H^n}$. Then X has non-empty interior, and $X \cap H^n$ is a manifold with boundary homeomorphic to $\partial H^n \setminus X$ and interior homeomorphic to an open (Euclidean) ball.*

Proof. Since X is full, there are n points in the interior of X which are in general position. These points form the vertices of an n-simplex and such a simplex has non-empty interior by Lemma 5.1.6.

Pick an arbitrary $x_0 \in \text{int}\, X$, and let U be an open ball about x_0. For $v \in \partial H^n$, let R_v be the hyperbolic ray from x_0 to v. This ray meets ∂X in exactly one point, and the function sending v to this point is a homeomorphism from $\partial H^n \setminus X$ to $\partial X \cap H^n$; in a similar way we define a homeomorphism from H^n to int $X \cap H^n$. ⅅ

Recall, a path $\alpha : [0,1] \to X$ in a metric space (X, ρ) is **rectifiable** if the limit

$$\inf \sum_{i=1}^{n} d(\alpha(\xi_{i-1}, \xi_i))$$

(where the infimum is taken over, for all $n \in \mathbb{N}$, the set of finite sequences $(\xi_0, \dots, \xi_n)$ where $\xi_0 = 0$ and $\xi_n = 1$) exists and is finite.

We see that $\partial X \cap H^n$ is a manifold-with-boundary, and so the connected components of the boundary are path-connected. Further, each pair of points may be joined by a rectifiable path.

5.1.8 Lemma. *Let X be a closed convex subset of $\overline{H^n}$; let x and y be points in the same component of $\partial X \cap H^n$. Then there is a rectifiable path joining x and y.*

Proof. Pick a point z in the interior of X. Let $R_x : \mathbb{R}_{\geq 0} \to \overline{H^n}$ be the ray from z through x such that $R_x(1) = x$ and $\lim_{t\to\infty} R_x(t)$ is on ∂H^n; similarly define a ray R_y from z through y such that $R_y(1) = y$. Observe now that $R_x(37)$ and $R_y(37)$ lie in $H^n \setminus X$ and retract to x and y respectively. Both of these lie in the same connected open subset of H^n so there is a rectifiable path joining them; since r is distance-reducing this path retracts to a rectifiable path in $\partial X \cap H^n$. ∎

Remark. The point of the proof is that it is a standard fact that we may find rectifiable paths in *open* path-connected subsets of H^n, but in arbitrary non-open path-connected subsets it might be the case that the paths between two points are wild. (Example?)

We may therefore define a metric d_S on each component S of $\partial X \cap H^n$: namely, send (x, y) to the infimum of the lengths of rectifiable paths from x to y.

Next we define the notion of 'open neighbourhood' which will be useful in the sequel.

5.1.9 Definition. Let x, v be distinct points in $\overline{H^n}$ and let $\varepsilon, \delta > 0$. The **open shell** $\mathrm{Sh}(x, v, \varepsilon, \delta)$ is the intersection of

$$\{y \in \overline{H^n} : |d(v, y) - d(v, x)| < \varepsilon\}$$

— that is, the two-sided ε-neighbourhood of the sphere centred at v through x — and the hyperbolic cone with vertex v, axis the ray from v to x, and vertex angle δ. We call x the **centre** of the shell, and v the **vertex** of the shell.

5.2 Foliations and laminations

Let Λ be a closed subset of S^{n-1}, and consider the convex hull $\mathrm{h{-}conv}\,\Lambda \cap H^n$ (see Fig. 5.3). The edges of this convex hull are geodesics in H^n which are eventually parallel and tend to (missing) points on S^{n-1}. The 'flat pieces' between the edges are curved inwards, and so we obtain a structure not unlike that of folded — more precisely, *pleated* — papers. In addition, the patterns look like they are obtained by gluing layers of wood back to back (the edges being the glue layers and the flat pieces being the wood panels end-on), and so we call them *lamination patterns*. In this section, we make these notions precise.

A **pseudogroup** on a topological space X is a system $\mathscr{G}$ of homeomorphisms between open sets of X such that

1. $\{\mathrm{dom}\, f : f \in \mathscr{G}\}$ is an open cover of X;
2. $\mathscr{G}$ is closed under restriction to open subsets of domains;
3. $\mathscr{G}$ is closed under compositions whenever they are defined;

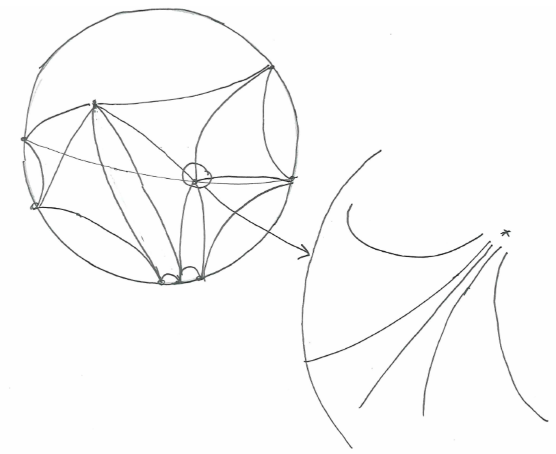


Figure 5.3: The convex hull boundary of a closed subset of S^{n-1}.

4. if $f : U \to V$ is a homeomorphism between $U, V \subseteq X$ open, and U is covered by open sets $\{U_\alpha\}$ such that for each α the restriction f_{U_α} lies in $\mathscr{G}$, then f itself lies in $\mathscr{G}$.

(Observe that this is not unlike the definition of a sheaf on X.)

5.2.1 Definition. If $\mathscr{G}$ is a pseudogroup on $\mathbb{R}^n$, then a topological space X is a $\mathscr{G}$**-manifold** if X is a manifold which admits an atlas whose transition maps lie in $\mathscr{G}$.

5.2.2 Example.

1. If $\mathscr{Trivial}$ is the set $\{1_U : U \text{ open in } X\}$ then an $\mathscr{Trivial}$-manifold is a set of discrete points.

2. If $\mathscr{Top}$ is the set of all homeomorphisms between open sets of $\mathbb{R}^n$ then the notion of a $\mathscr{Top}$-manifold is equivalent to the usual definition of topological manifold.

3. If $\mathscr{C}^r$ is the set of all homeomorphisms between open sets of $\mathbb{R}^n$ which are of class C^r, then we obtain $\mathscr{C}^r$-manifolds (differentiable manifolds of class C^r); if $r = \infty$ we obtain smooth manifolds; and if $r = \omega$ (i.e. we allow only analytic maps) then we obtain real analytic manifolds.

4. Let $\mathscr{Conf}$ be the set of all homeomorphisms between open sets of $\mathbb{R}^2$ which are also biholomorphic maps with the induced complex structure obtained by identifying $\mathbb{R}^2$ with $\mathbb{C}$ in the usual way; then $\mathscr{Conf}$-manifolds are Riemann surfaces.

Write $\mathbb{R}^n = \mathbb{R}^{n-k} \times \mathbb{R}^k$, and let $\mathscr{G}$ be a pseudogroup generated by diffeomorphisms between open subsets of $\mathbb{R}^n$ with the property that the second k components do not depend on the first k: that is, elements of $\mathscr{G}$ are homeomorphisms $\phi : \mathbb{R}^{n-k} \times \mathbb{R}^k \supseteq U \to V \subseteq \mathbb{R}^{n-k} \times \mathbb{R}^k$ such that there exist $\phi_1 : U \to V \cap \mathbb{R}^{n-k}$ and $\phi_2 : U \cap \mathbb{R}^k \to V \cap \mathbb{R}^k$ such that

$$\phi(x, y) = (\phi_1(x, y), \phi_2(y)) \tag{5.2.3}$$

for all $x \in \mathbb{R}^{n-k}, y \in \mathbb{R}^k$ (so ϕ 'takes horizontal factors to horizontal factors'). Equivalently, the Jacobian of ϕ always takes block form with the zero matrix in the lower-left quadrant.

A $\mathscr{G}$-structure on a manifold M is called a **foliation** on M. A **leaf** is the inverse image, for some $y \in \mathbb{R}^k$, of $\mathbb{R}^{n-k} \times \{y\}$ under a $\mathscr{G}$-chart on M.

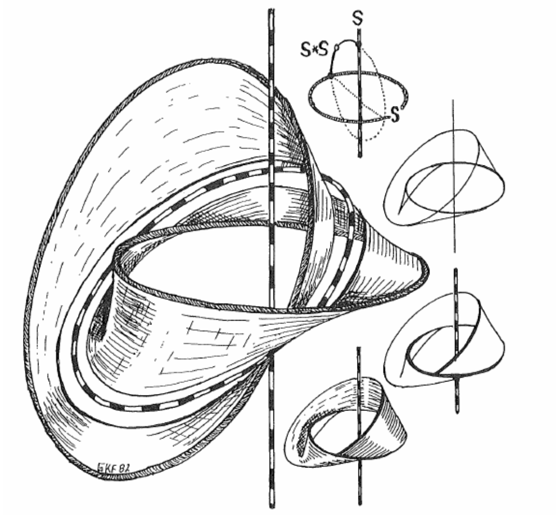


Figure 5.4: The Hopf fibration of S^3. Figure from [21, p. 175].

5.2.4 Example. The trivial dimension-k foliation of $\mathbb{R}^n$ is the $\mathcal{G}$-structure on $\mathbb{R}^n$ such that $\mathcal{G}$ is the largest pseudogroup satisfying the foliation conditions. The leaves are the additive cosets of $\mathbb{R}^k$ in $\mathbb{R}^n$.

5.2.5 Example. A codimension-1 foliation of the figure 8 knot complement is depicted as Fig. 6.2 above.

5.2.6 Example. The famous **Hopf fibration** is a codimension-1 foliation of the 3-sphere; see Fig. 5.4 and a codimension-2 version at `https://www.youtube.com/watch?v=AKotMPGFJYk`.

5.2.7 Example. The **Reeb foliation** is a codimension-1 foliation of the 3-sphere; see Fig. 5.5.

To round off the discussion of foliations, we list some major results in the field.

- A manifold M has a codimension-1 foliation iff $\chi(M) = 0$ (Thurston, 1976);
- No foliation of S^3 by surfaces is real analytic (Haefliger, 1958);
- Every codimension-1 foliation of S^3 has a leaf which is a torus (Novikov, 1965).

5.2.8 Definition. A **lamination** L on a manifold M is a closed subset $A \subseteq M$ (the **support** of L) together with a local product structure on A: that is, there is a family of open sets $\{U_i\}$ of M which cover A, together with charts $\phi_i : U_i \to \mathbb{R}^{n-k} \times \mathbb{R}^k$ for each i, such that $\phi_i(A \cap U_i) = \mathbb{R}^{n-k} \times B$ ($B \subseteq \mathbb{R}^k$) for each i, and such that the transition maps are of the form

$$\phi_i \phi_j^{-1}(x, y) = (f_{i,j}(x, y), g_{i,j}(y)) \tag{5.2.9}$$

for all i, j and for all $x \in \mathbb{R}^{n-k}$, $y \in B$.

Comparing Eq. (5.2.9) to Eq. (5.2.3), we see that a lamination is, in some sense, a foliation of a closed subset of M. We define **leaves** analogously, as 'horizontal slices' of A.

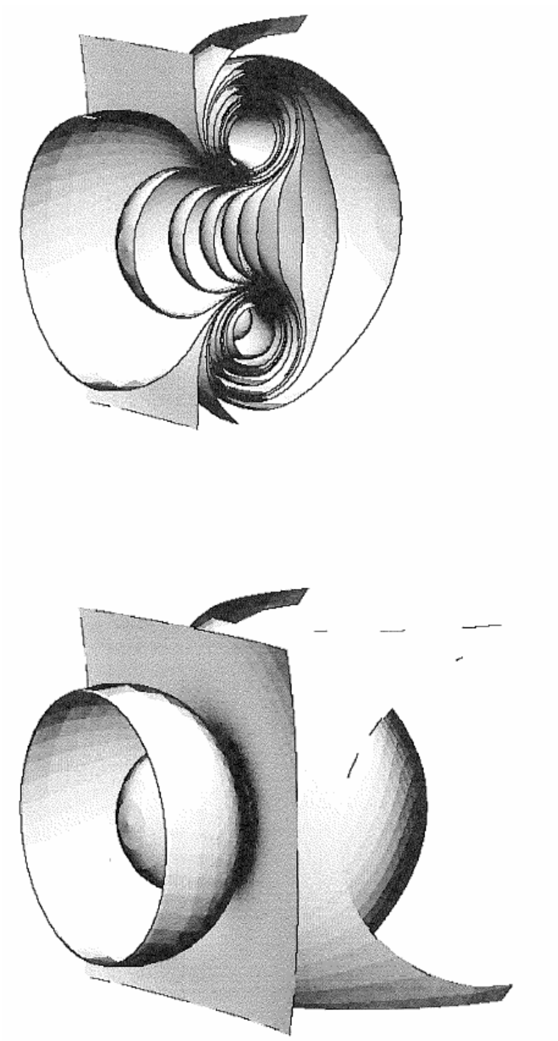


Figure 5.5: The Reeb fibration of S^3. Figure from https://www.ms.u-tokyo.ac.jp/~tsuboi/showroom/public_html/animations/gif/stereograph40/stereograph4034i.html.

5.2.10 Example. Let $K \subseteq M$ be a closed subset of a manifold M with foliation; then the closure of the set of leaves of the foliation which intersect with K is a lamination on M,

5.2.11 Example. The closure of any submanifold is a lamination with a single leaf.

A **geodesic lamination** is a lamination on a hyperbolic surface in which each leaf is a geodesic arc.

5.2.12 Lemma. *A geodesic lamination on a surface is precisely a disjoint union of geodesics on the surface.*

5.3 Pleated surfaces

Let M be a hyperbolic 3-manifold. A **pleated surface** in M is a complete hyperbolic surface S together with an isometric map $f : S \to M$ (to be more precise, the map f is isometric onto the intrinsic metric of the image $f(S)$, not isometric with respect to the ambient metric of M — this is clearly the property we want, if we think of f as a map which 'folds up' S: we wish distances to be preserved if we walk along S, but in general the distances if we allow ourselves to move off $f(S)$ and through M will be smaller than the distances we walk if we restrict ourselves to $f(S)$) such that every point $s \in S$ lies in the interior of a geodesic in S which is mapped by f to a geodesic arc in M, and such that f is homotopically incompressible (that is, $f_* : \pi_1(S) \to \pi_1(M)$ has trivial kernel — the embedding does not trivialise loops even if we allow deformation through M).

Given such a pleated surface (S, f), the **pleating locus** is defined to be the set of points $s \in S$ such that only a single geodesic through s which is mapped to a geodesic arc with respect to the metric of M. The connected components of the complement of the pleating locus are called the **flat pieces** of the pleating.

Recall that a map $f : X \to Y$ of Riemann manifolds is **totally geodesic** if, for all geodesics $\alpha \subseteq X$, the image $f(\alpha)$ is a geodesic in Y.

The following lemma is geometrically evident.

5.3.1 Lemma. *If (S, f) is a pleated surface in M, then the pleating locus of the system is a geodesic lamination in M.*

Proof. We shall write d_S for the intrinsic metric on S and $f(S)$, and d_M for the metric on M. Hence $d_S(x, y) = d_S(f(x), f(y)) \geq d_M(f(x), f(y))$ for all $x, y \in S$.

Let γ be the pleating locus of the system. If $x \notin \gamma$ then there exist two transverse geodesics through x; pick a and b points on one and c and d points on the other, such that $a < x < b$ and $c < x < d$ with respect to an orientation of each geodesic and such that the geodesic arcs $[a, b]$ and $[c, d]$ lie in the same flat piece Π as x. Let θ be the angle at x of the triangle axc; then the angle at x of bxc is $\pi - \theta$. Thus, in S, we have the triangle indicated in Fig. 5.6. Since $d_M(f(a), f(x)) = d_S(a, x)$, $d_M(f(c), f(x)) = d_S(c, x)$, and $d_M(f(a), f(x)) \leq d_S(a, x)$, the angle at $f(x)$ of the triangle $f(a)f(x)f(c)$ is at most θ. By a similar argument, the angle at $f(x)$ of $f(c)f(x)f(b)$ is at most $\pi - \theta$. But these angles must add to π, and so both inequalities must in fact be equalities; that is, f preserves angles at x, thus $d_M(f(a), f(x))$ in fact *equals* $d_S(a, x)$, and so f maps the geodesic $[a, c]$ to the geodesic $[f(a), f(c)]_M$ (the subscript denoting that this is a geodesic with respect to the metric of M): so f is totally geodesic in a neighbourhood of x in S; and this neighbourhood must therefore lie entirely in Π, so Π is open and γ is closed.

It remains only to show that if $x \in \gamma$ then the (unique, by definition) geodesic arc containing x which is mapped to a geodesic in M by f is in fact a complete geodesic. Let α be the maximal geodesic arc containing x with the property that $f\alpha$ is a geodesic in M; we show that α is complete. If not, α

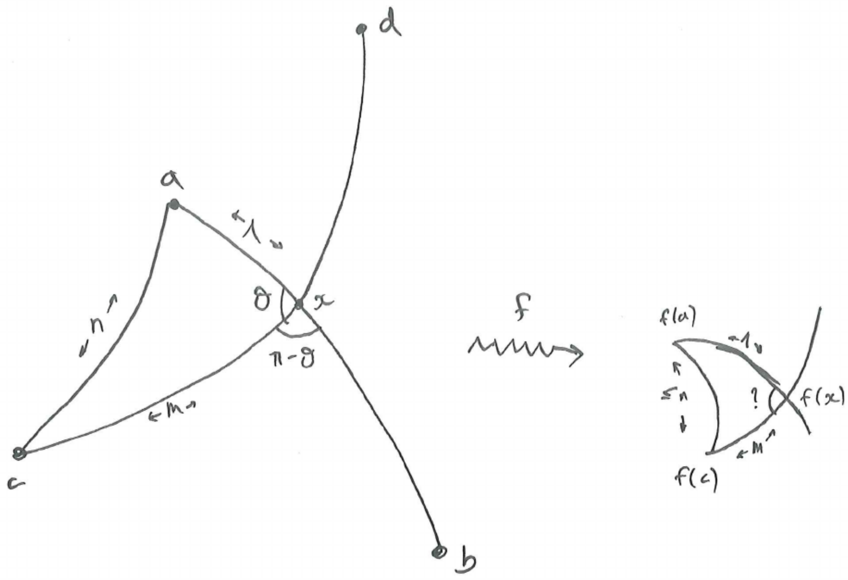


Figure 5.6: Action of f on flat pieces in the proof of Lemma 5.3.1.

has an endpoint a; let b be a point on α on the opposite side of x to a. Observe that $a \in \gamma$ (otherwise, $S \setminus \gamma$ could not be open) and so there is some geodesic arc containing a in its interior which is mapped by f to a geodesic arc in M; say this arc is $[c, d]$. Observe that $[c, d] \cap \alpha = \{a\}$, otherwise a portion of $[c, d]$ could be glued onto α to form a geodesic strictly containing α. This situation is depicted in Fig. 5.7. By a very similar argument to the paragraph on flat pieces, applied to the triangles acb and adb in that figure, we see that f maps $[b, c]$ to the geodesic $[f(b), f(c)]_M$ and maps $[b, d]$ to $[f(b), f(d)]_M$. Hence f acts to preserve geodesics on each side of acb and adb; thus in particular f preserves the geodesic $[c, d]$ which contradicts uniqueness of a preserved geodesic through a.

5.3.2 Example. The geodesics on the 2-torus depicted in Fig. 5.8 form a lamination. If they cover the whole 2-torus, then the lamination is in fact a foliation.

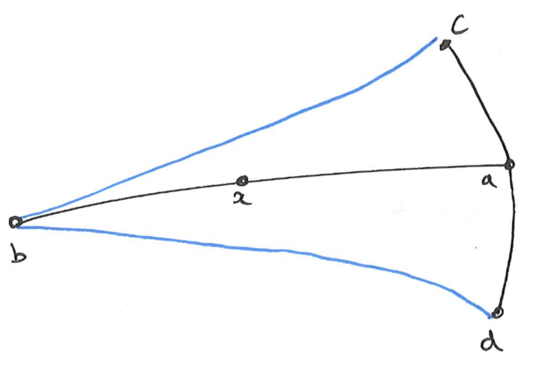


Figure 5.7: Action of f on geodesics in the proof of Lemma 5.3.1.

Figure 5.8: A geodesic lamination on the 2-torus.

5.4 Roofs, and their lowering

In this section, we prove various results on finite approximations to convex hulls: we follow [18, sections 1.8 to 1.10].

5.4.1 Lemma. *Suppose P_1, P_2, and P_3 are hyperbolic planes without a common point of intersection that pairwise intersect transversely. Let C_1, C_2, and C_3 be the respective intersections of the planes with the sphere at infinity. Then, either*

1. *the C_i have no common intersection point: in this case, there is a unique circle C^* orthogonal to each of the C_i which determines a plane P^* orthogonal to all the P_i, and the three lines $P^* \cap P_i$ determine a triangle in P^* such that the vertex angles of the triangle are the dihedral angles between the various P_i; or*

2. *the C_i have a common intersection point ζ: in this case, if σ is any horosphere based at ζ, then the three curves $\sigma \cap P_i$ determine a Euclidean triangle in σ with vertex angles the dihedral angles between the various P_i.*

Proof. [18, lemma 1.10.1]

5.5 Pleating properties of the convex hull

Let Λ be a closed subset of S^{n-1}; for convenience we write S for $\partial\,\text{h-conv}\,\Lambda \cap H^n$.

Our main result (which is theorem 1.12.1 of [11]) is the following.

5.5.1 Theorem. *The boundary S is a pleated surface, with pleating locus the set of edges of* $\text{h-conv}\,\Lambda \cap H^n$ *(until we have proved that these indeed form a pleating locus, we will call these edges the* ***bending lines*** *of the surface).*

Proof. We show that S is a complete hyperbolic surface. We begin by showing that for each $x \in S$ there is an open neighbourhood of x isometric to an open set in H^2. If x lies in a flat piece of S, this is trivial. Assume therefore that x lies on a bending line l of S. Let U be a shell in $\overline{H^n}$ centred at x; we show that $U' := U \cap S$ may be mapped isometrically into H^2.

Fix orientations of l, H^2, and H^3. As noted above we may orient S such that the 'positive' side contains the interior of $\text{h-conv}\,\Lambda$. Fix also an isometric embedding $g : l \to H^2$, and let l' be the image of l under this embedding with the induced orientation. We will define a map $g : U' \to H^2$ which will extend this embedding and 'unfold' the bending line l while preserving orientation.

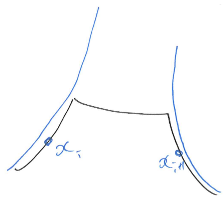


Figure 5.9: A bad choice of points for a polygonal approximation.

Definition of g. Let $y_1, y_2 \in l \cap U$ be distinct points. Given $z \in U'$, define gz to be the unique point in H^2 which lies on the correct side of l' and such that

$$d(gz, gy_1) = d_S(z, y_1), \text{ and} d(gz, gy_2) = d_S(z, y_2).$$

Checking that g is an isometry. The proof that g is an isometry is technical: the idea is to construct a sequence S_n of approximations to S about l which have finitely many facets; letting g_n be the embedding constructed similarly to g but for S_n rather than S, we see that each g_n is an isometry and so since $g_n \to g$ we have that g is an isometry. (One additional technical point is that the intrinsic metric is equivalent to the restriction of the H^3 metric, so this argument works.)

Checking that S is complete. If (x_n) is Cauchy with respect to the metric d_S, then (x_n) is also Cauchy with respect to the hyperbolic metric in H^3 and hence converges; since S is closed in H^3, the limit of (x_n) with respect to the hyperbolic metric lies in S; since the hyperbolic and intrinsic metrics are equivalent, we are done.

We place a transverse measure on the pleating locus of S: in this context, a **transverse measure** on the pleating locus is a regular measure (see Appendix B) defined on the set of embedded intervals in S which are transverse to every bending line that they meet.

For $x \in S$, let $\pi(x)$ be the set of oriented supporting hyperplanes at x, and let

$$Z(S) = \{(x, P(x)) : x \in S, P(x) \in \pi(x)\};$$

there is a natural topology on $Z(S)$, namely that induced as a subspace of $H^3 \times \mathbb{G}_2(H^3)$.

Any path in $Z(S)$ projcts to a path in S; conversely, if ω is a path on S then we may extend it to a path in $Z(S)$ (namely, lift $x \in \omega$ to a path joining (x, P) to (x, Q) where P and Q are the extreme supporting hyperplanes at x).

Suppose that $\omega : [0,1] \to Z$ is such a path; a **polygonal approximation** to ω is a choice of a partition $0 = t_0 < t_1 < \cdots < t_n = 1$ such that, if P_i denotes the second component of $\omega(t_i)$, $P_i \cap P_{i+1} \neq \emptyset$ for each i (this to avoid badly-behaved approximations like that of Fig. 5.9). We denote such an approximation by the sequence $(\omega(t_i) = (x_i, P_i))_{i=0}^n$ of pairs, so a polygonal approximation to a curve in S is precisely a choice of finitely many points on the curve and a finite non-empty set of supporting hyperplanes at x_i for each i such that each point of the curve has a roof over it.

For each i, let θ_i be the absolute value of the angle between the outward normals of P_{i-1} and P_i. Then the **bending measure** $\beta(\omega)$ is defined to be

$$\beta(\omega) := \inf \sum_{i=1}^{n} \theta_i$$

where the infimum is taken over all polygonal approximations to ω.

That β is indeed a measure is proved in [18, theorem 1.11.3]:

5.5.2 Lemma. *The bending measure is indeed a transverse measure.*

Chapter 6

3-manifolds

In this chapter, a **Kleinian group** is a discrete subgroup of $\mathrm{Isom}^+(H^n)$.

6.1 Homotopy

We recall some standard facts from algebraic topology, see for instance [33, chapter 12].

Let X be a topological space; a continuous map $q : \hat{X} \to X$ is called a **covering map** if $\hat{X}$ is connected and locally path connected, and if for all $x \in X$ there exists a neighbourhood U of x such that $q^{-1}(U)$ is a disjoint union of connected open subsets of $\hat{X}$ called the **sheets** of the covering over U, with the property that each sheet is homeomorphically mapped onto U by q.

Given such a covering map, an **automorphism** or **deck transformation** of q is a homeomorphism $f : \hat{X} \to \hat{X}$ such that $qf = q$. The set of all automorphisms under composition forms a group, $\mathrm{Aut}_q \hat{X}$.

A covering map $q : \hat{X} \to X$ is **normal** or **regular** if the induced subgroup $q_* \pi_1(\hat{X}, x_0)$ is normal in $\pi_1(X, q(x_0))$ for some x_0. This is equivalent to the subgroup being normal for all choices of x_0 ([33, proposition 11.35]). The map q is said to have the **homotopy lifting property** if for all maps $f : Y \to \hat{X}$ and for all homotopies $\overline{f_t}$ of $\overline{f} = \pi f$ in X there exists a homotopy f_t of f in $\hat{X}$ lifting the homotopy $\overline{f_t}$.

An action of a group Γ on a space X by homeomorphisms is called **effective** if the homeomorphism $f_\gamma : X \to X$ for $\gamma \in \Gamma$ defined by $f_\gamma(x) = gx$ is the identity iff $\gamma = 1$; i.e. if $\gamma = 1 \iff \forall_{x \in X} \gamma x = x$.

6.1.1 Theorem. *Let X be a locally path-connected space and let Γ act on X effectively. Then the quotient $\pi : X \to X/\Gamma$ is a covering map iff the action is freely discontinuous. In this case, π is a normal covering map and* $\mathrm{Aut}_\pi(X) = \Gamma$.

Remark. Observe that in the surface case we studied previously, this theorem applies only to the quotient $°\Omega/\Gamma$ and not to the more general case $\mathcal{R}_\Gamma = \Omega/\Gamma$; the problem is that in the latter case we obtain branch points and so we get a branched cover and an orbifold.

We now apply this theory to some examples, following [43, section 8.1].

6.1.2 Lemma. *If Γ is a non-elementary Kleinian group and $\Gamma' \leq \Gamma$ is a nontrivial normal subgroup, then* $\Lambda(\Gamma') = \Lambda(\Gamma)$.

Proof. Observe that Γ' is infinite (suppose not; by normality, Γ leaves invariant $\mathrm{Fix}_{H^n}(\Gamma)$, in particular since Γ' is finite it is of elliptic type and so $\mathrm{Fix}_{H^n}(\Gamma)$ is non-empty and finite, thus Γ has a finite orbit

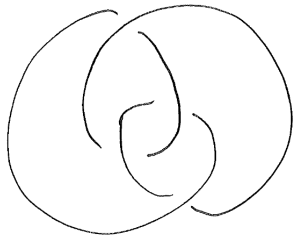


Figure 6.1: A figure 8 knot diagram.

in H^n and so is elementary which contradicts the hypotheses). Hence by compactness of $\overline{H^n}$ the limit set $\Lambda(\Gamma')$ is nontrivial. By normality, $\gamma\Lambda(\Gamma') \subseteq \Lambda(\Gamma')$, and hence by part 2 of Theorem 3.3.21 we have $\Lambda(\Gamma) \subseteq \Lambda(\Gamma')$. The opposite inclusion is trivial.

6.1.3 Example. Let M be a hyperbolic surface. Then $\pi_1(M)$ is a group of isometries of a plane in H^3: indeed, by the Klein-Poincaré theorem (Theorem 4.1.7) we have that $M \simeq \hat{M}/\Gamma$ for some Fuchsian group Γ and by Theorem 6.1.1 we have $\Gamma \simeq \pi_1(M)$. In this case, $\pi_1(M)$ has a limit set contained in a circle (it is Fuchsian).

6.1.4 Example. If M is a closed hyperbolic 3-manifold, then $\pi_1(M)$ is a Kleinian group with limit set $\hat{\mathbb{C}}$. This will follow from the fact which we shall prove later (Corollary 6.2.7) that, for every hyperbolic 3-manifold M, $M \simeq H^3/\Gamma$ for some Kleinian group Γ, and $\partial M \simeq \Omega(\Gamma)/\Gamma$; for the latter set to be empty we must have $\Omega(\Gamma) = \emptyset$ and thus $\Lambda(\Gamma) = \hat{\mathbb{C}}$.

Suppose that M is closed and fibred over the circle (i.e. there exists $q : M \to S^1$ a covering map with the homotopy lifting property). Intuitively, this means that M may be spanned by a union of 2-dimensional surfaces parameterised continuously by a point on the circle. In this case, the fundamental group of the fibres is a normal subgroup of $\pi_1(M)$ (why? does this follow from the algebraic topology?) and so by Lemma 6.1.2 has limit set equal to that of $\pi_1(M)$; by the previous paragraph, this is $\hat{\mathbb{C}} = S^2$.

6.1.5 Example. Let k be the figure 8 knot — that is, the knot with diagram depicted in Fig. 6.1 — and let $M = \hat{\mathbb{R}}^3 \setminus k$. We find, using standard techniques, that

$$\pi_1(M) = \langle A, B \,:\, ABA^{-1}BA = BAB^{-1}AB \rangle$$

(see, e.g., [9, example 3.8]). This manifold fibres over S^1: the fibres are the Seifert surfaces of the knot (see Fig. 6.2), which may be shown to have genus 1 (though this is not intuitively clear, see [1]) and thus by standard surface theory each fibre is a punctured torus, F. A computation shows that $\pi_1(F) = \langle AB^{-1}, A^{-1}B \rangle$ (see, e.g. [9, theorem 4.6]). By the previous example, the limit set is S^2.

6.2 Developing maps and holonomy

We follow [44, section 3.4] and [38, section 8.4].

Let X be a connected real analytic manifold (that is, X is an $\mathscr{C}^\omega$-manifold as described in Example 5.2.2), let G be a group of real analytic diffeomorphisms on X, and let M be a (G, X)-manifold.

Fix a point $x_0 \in M$, and let $\gamma : [0,1] \to M$ be a curve in M with $\gamma(0) = x_0$. The image of an initial segment of γ may be lifted to X via any chart around x_0; our goal is the 'extension' of this lifting along the length of X. To do this, we will pick a chain of charts $(U_0, \phi_0), \ldots, (U_{n-1}, \phi_{n-1})$ along γ and then

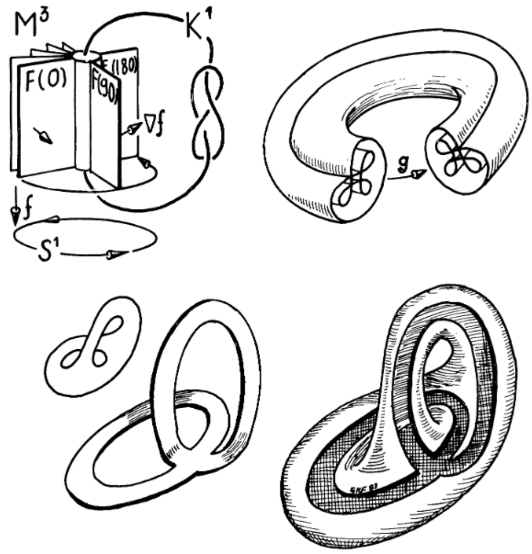


Figure 6.2: A fibration of the knot complement of the figure 8 knot. Figure from [21, p. 159].

'adjust' each so that their images in X match up (Fig. 6.3a). We will need to do a bit of work to ensure this is well-defined.

Since the image of γ is compact, there exists a partition $0 = t_0 < t_1 < \cdots < t_n = 1$ of $[0, 1]$ such that, for each $0 \leq i < n$, the image $\gamma[t_i, t_{i+1}]$ lies within a single chart, (U_i, ϕ_i). For each such i there is an element $g_i \in G$ which locally agrees with the transition map

$$\phi_i \phi_{i+1}^{-1} : \phi_{i+1}(U_{i+1} \cap U_i) \to \phi_i(U_{i+1} \cap U_i);$$

as usual, in what follows we will identify the transition maps and the group elements whenever we are working locally in a chart.

Observe that $\phi_i \gamma \restriction_{[t_i, t_{i+1}]}$ and $g_i \phi_{i+1} \gamma \restriction_{[t_{i+1}, t_{i+2}]}$ are curves in X, and that they agree at t_{i+1}:

$$g_i \phi_{i+1} \gamma \restriction_{[t_{i+1}, t_{i+2}]}(t_{i+1}) = \phi_i \phi_{i+1}^{-1} \phi_{i+1} \gamma \restriction_{[t_{i+1}, t_{i+2}]}(t_{i+1}) = \phi_i \gamma \restriction_{[t_{i+1}, t_{i+2}]}(t_{i+1}).$$

In particular, we may glue these curves together in X to form a continuous curve $\hat{\gamma} : [0, 1] \to X$ by the rule

$$\hat{\gamma}(t) := \begin{cases} \phi_0 \gamma \restriction_{[t_0, t_1]} & t \in [t_0, t_1] \\ g_0 \phi_1 \gamma \restriction_{[t_1, t_2]} & t \in [t_1, t_2] \\ \vdots & \\ g_0 \cdots g_{n-2} \phi_{n-1} \gamma \restriction_{[t_{n-1}, t_n]} & t \in [t_{n-1}, t_n]. \end{cases}$$

The curve $\hat{\gamma}$ is called the **analytic continuation** of $\phi_0 \gamma$ along γ, and the chart $g_0 \cdots g_{n-2} \phi_{n-1}$ is called the analytic continuation of ϕ_0 along γ; *a priori* the result depends on the choice of charts ϕ_i about each segment of the partition of $[0, 1]$, and on the choice of the partition (t_i). The following lemmata show that there is no such dependence.

6.2.1 Lemma. *If $\gamma : [0, 1] \to M$ is a curve in M with $\gamma(0) = x_0$, a partition $0 = t_0 < t_1 < \cdots < t_n = 1$ of $[0, 1]$ is chosen as above, and for each i there exists a pair of charts (U_i, ϕ_i) and (V_i, ψ_i) with $U_0 = V_0$ and $\phi_0 = \psi_0$ such that $\gamma[t_i, t_{i+1}]$ is contained in both $\phi_i(U_i)$ and $\psi_i(V_i)$, and $\hat{\gamma}_\phi$ and $\hat{\gamma}_\psi$ are the analytic continuations obtained with respect to the respective charts, then $\hat{\gamma}_\phi = \hat{\gamma}_\psi$.*

Proof. For all i, let $g_i := \phi_i\phi_{i+1}^{-1}$ and let $h_i := \psi_i\psi_{i+1}^{-1}$. Given i, both U_i and V_i contain $\gamma[t_i, t_{i+1}]$ and hence it suffices to show that, on $U_i \cap V_i$,

$$g_0 \cdots g_{i-1}\phi_i = h_0 \cdots h_{i-1}\psi_i.$$

We proceed by induction on i. The case $i = 0$ is by hypothesis; suppose that $g_0 \cdots g_{i-2}\phi_{i-1} = h_0 \cdots h_{i-2}\psi_{i-1}$.. The inductive step will follow from rigidity of real analytic functions, which we state as a lemma:

Lemma. *If $f, g : U \to \mathbb{R}^m$ (U open in $\mathbb{R}^n$) are analytic and agree on an open subset of U, then they agree everywhere.*

Indeed, let $f_i \in G$ be the group element locally agreeing on $\phi_i(U_i \cap V_i)$ with the 'automorphic transition' $\psi_i\phi_i^{-1}$. By inductive assumption, this transition map equals

$$\psi_i(h_0 \cdots h_{i-2}\psi_{i-1})^{-1}(g_0 \cdots g_{i-2}\phi_{i-1})\phi_i^{-1} = \psi_i\psi_{i-1}^{-1}h_{i-2}^{-1} \cdots h_0^{-1}g_0 \cdots g_{i-2}\phi_{i-1}\phi_i$$

on $\phi_i(U_{i-1} \cap U_i \cap V_{i-1} \cap V_i)$. On the same set we have the equality

$$(\psi_i\psi_{i-1}^{-1})(h_{i-1}^{-1} \cdots h_0^{-1})(g_0 \cdots g_{i-1})(\phi_{i-1}\phi_i^{-1}) = h_{i-1}^{-1} \cdots h_0^{-1}g_0 \cdots g_{i-1}$$

by definition of g_{i-1} and h_{i-1}. In particular, f_i locally agrees with

$$\psi_i\psi_{i-1}^{-1}h_{i-2}^{-1} \cdots h_0^{-1}g_0 \cdots g_{i-2}\phi_{i-1}\phi_i = h_{i-1}^{-1} \cdots h_0^{-1}g_0 \cdots g_{i-1}.$$

Hence by the internal lemma above, $f_i = h_i^{-1} \cdots h_0^{-1}g_0 \cdots g_i$ on $U_i \cap V_i$. In particular, in $U_i \cap V_i$ we have

$$g_0 \cdots g_{i-1}\phi_i = (h_0 \cdots h_{i-1})(h_0 \cdots h_{i-1})^{-1}g_0 \cdots g_{i-1}\phi_i = (h_0 \cdots h_{i-1})f_i\phi_i = (h_0 \cdots h_{i-1})\psi_i$$

as was to be shown.

6.2.2 Lemma. *If $\gamma : [0,1] \to M$ is a curve in M with $\gamma(0) = x_0$, and partitions $0 = t_0 < t_1 < \cdots < t_n = 1$ and $0 = s_0 < s_1 < \cdots < s_n = 1$ of $[0,1]$ are chosen as above, and $\hat{\gamma}_t$ and $\hat{\gamma}_s$ are the analytic continuations obtained with respect to the respective partitions, then $\hat{\gamma}_s = \hat{\gamma}_t$.*

Proof. Pass to the partition $\{t_i\} \cup \{s_i\}$; this partition must give an analytic continuation equal to both $\hat{\gamma}_t$ and $\hat{\gamma}_s$.

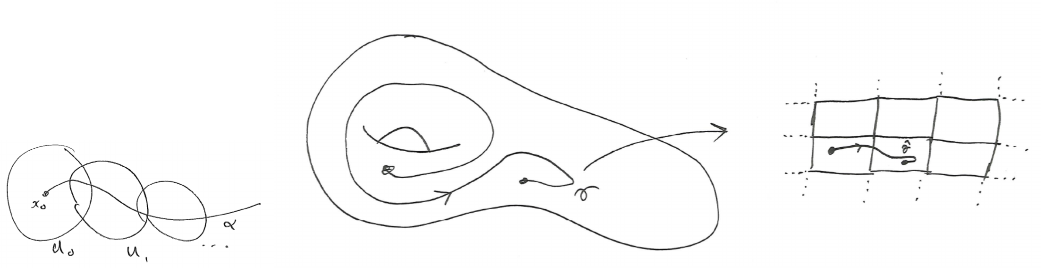
Finally, we show that analytic continuation is independent of homotopy.

6.2.3 Theorem. *Let $\gamma, \eta : [0,1] \to M$ be curves in M such that $\gamma(0) = \eta(0)$ and $\gamma(1) = \eta(1)$. If γ and η are homotopic with fixed endpoints in M, then $\hat{\gamma}$ and $\hat{\eta}$ are homotopic with fixed endpoints in X.*

Proof.

Recall that we may view the universal cover $\tilde{M}$ of the manifold M as the quotient of the space of all curves $\gamma : [0,1] \to M$ with $\gamma(0) = x_0$ (where x_0 is a fixed basepoint on M) by homotopy equivalence with fixed endpoints; the projection of a homotopy class $[\gamma]$ is $\pi([\gamma]) := \gamma(1)$.

For fixed basepoint x_0 and fixed chart ϕ_0 about x_0, the **developing map** is the map $D : \tilde{M} \to X$ which, around each $[\gamma] \in \tilde{M}$, agrees locally with the germ of the analytic continuation of ϕ_0 along γ. This is locally a (G, X)-diffeomorphism.



(a) Charts along a curve. (b) A curve on a torus and its analytic continuation in $\mathbb{R}^2$.

Figure 6.3: Analytic continuation along a curve.

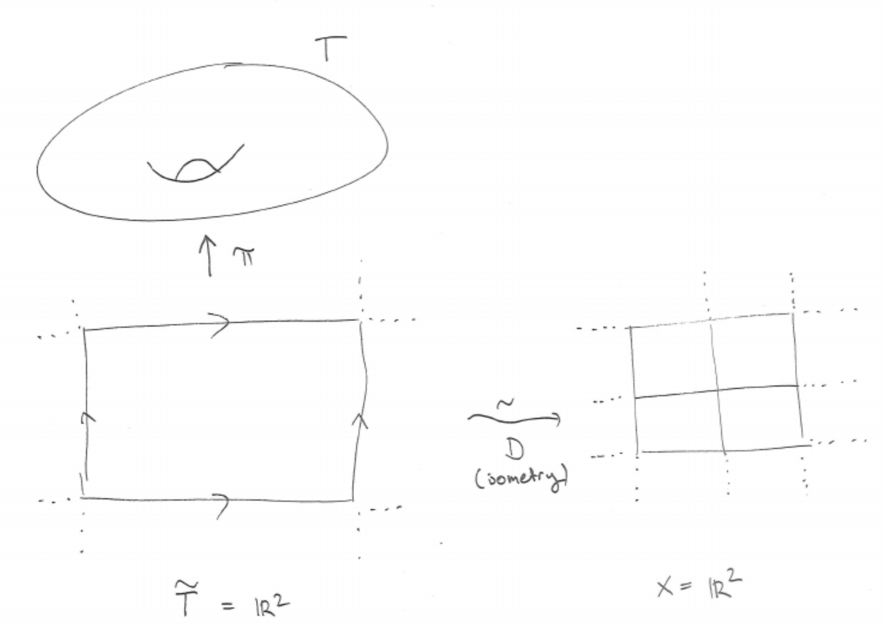


Figure 6.4: Euclidean torus development (Example 6.2.4).

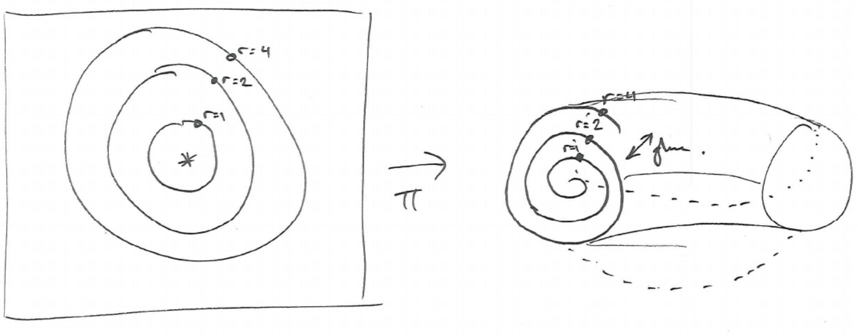


Figure 6.5: The torus as an affine quotient of $\mathbb{R}^2 \setminus \{0\}$.

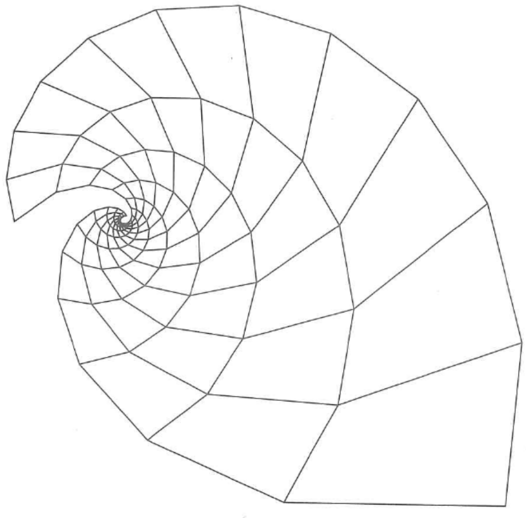


Figure 6.6: Affine torus development (Example 6.2.5). Figure from [44, figure 3.17].

6.2.4 Example. Consider the quotient $\mathbb{R}^2/\Gamma$ where Γ is the group of affine transformations generated by $(x, y) \mapsto (x + 1, y)$ and $(x, y) \mapsto (x, y + 1)$.This quotient is homeomorphic to the torus T (Fig. 6.4) and so T may be given the structure of an (Isom($\mathbb{R}^2$), $\mathbb{R}^2$)-manifold. Picking a point $x_0 \in \mathbb{R}^2$ (which is the universal cover for T) we see that the developing map sends the coordinate rectangles in $\mathbb{R}^2$ to themselves: the development is the standard rectangular tiling of $\mathbb{R}^2$.

6.2.5 Example. Consider the quotient $(\mathbb{R}^2 \setminus \{0\})/\Gamma$ where Γ is the group of affine transformations generated by $x \mapsto 2x$. This quotient is homeomorphic to the torus T (Fig. 6.5) and so T may be given the structure of an $(\mathrm{Aff}(\mathbb{R}^2), \mathbb{R}^2)$-manifold. Picking a point $x_0 \in \mathbb{R}^2$ (which is the universal cover for T) and noting that the universal covering map $\pi : \mathbb{R}^2 \to \mathbb{R}^2 \setminus \{0\}$ is given by $x + yi \mapsto \exp(x + yi)$, we see that the developing map sends the coordinate rectangles in $\mathbb{R}^2$ to the 'logarithmic spiral' of Fig. 6.6.

Let α now be a loop on M, with basepoint x_0. Pick ϕ_0 a chart around x_0; we may use analytic continuation to find a second chart ϕ_0^α about x_0. Since ϕ_0^α and ϕ_0 are compatible charts, there is an element $g_\alpha \in G$ with the transition property $\phi_0^\alpha = g_\alpha \phi_0$. If $T_\alpha : \pi_1(M, x_0) \to \pi_1(M, x_0)$ is the map $\beta \mapsto \alpha\beta$ then we have a commutative square $DT_\alpha = g_\alpha D$.

We now have a homomorphism $H : \pi_1(M) \to G$ sending $\alpha \mapsto g_\alpha$; this is the **holonomy map** of M (which depends on x_0, but only up to conjugacy in G). The image $H(\pi_1(M))$ is the **holonomy group**.

We say that M is a **complete** (G, X)-manifold if the developing map is a covering map. If M is complete and X is simply connected, then the map D sets up an identification $\tilde{M} = X$; in fact:

6.2.6 Theorem. *If G is a group of analytic diffeomorphisms on a simply connected space X, and if M is a complete (G, X)-manifold, then there is a canonical (up to G-translation) identification $M \simeq_{(G,X)\text{-diffeo.}} X/\Gamma$, where Γ is the holonomy group of M.*

Proof. Ratcliffe 8.5.9 ⅁

6.2.7 Corollary. *Let M be a complete hyperbolic manifold; then $M \simeq_{\text{hyp. diffeo.}} X/\Gamma$ for some Kleinian group Γ.* ⅁

Remark. The above theory can be extended to orbifolds: the analogen of Theorem 6.2.6 is [38, theorem 13.3.10].

6.3 Hyperbolic convexity

We follow [43, section 8.3]

6.3.1 Definition. A complete hyperbolic manifold with boundary M is **convex** if each path in M is homotopic to a geodesic arc.

6.3.2 Example. For submanifolds of H^n, this is equivalent to the usual definition of convexity.

6.3.3 Proposition. *The manifold M is convex iff the developing map $D : \tilde{M} \to H^n$ is a homeomorphism onto a convex subset of H^n.*

Proof. Suppose that $D(\tilde{M})$ is convex. If α is a path in M, then α lifts to a path $\tilde{\alpha}$ in $\tilde{M}$; by convexity of $D(\tilde{M}) \simeq \tilde{M}$, $\tilde{\alpha}$ is homotopic to a geodesic arc in $\tilde{M}$; then the projection of this arc is a geodesic arc in M homotopic to α.

Conversely, suppose that M is convex. Recall that D is a local homeomorphism, so it remains to show that D is injective and that $D(\tilde{M})$ is convex. Let $x, y \in \tilde{M}$. There exists a path joining πx and πy in M which (by convexity) is homotopic to a geodesic arc; by the lifting property; this path lifts to a geodesic in $\tilde{M}$ joining x and y; and D sends distinct endpoints of a geodesic to distinct points in H^n. Finally suppose α is a path in $D(\tilde{M})$; then there is a lift of α in $\tilde{M}$ which is homotopic to a geodesic; and the image of this geodesic under D remains a geodesic.

We say that M is **locally convex** if each $x \in M$ has a neighbourhood isometric to a convex subset of H^n. Clearly convexity implies local convexity. The converse is also true, but is less trivial:

6.3.4 Theorem. *If M is locally convex, then M is convex.*

Proof. Omitted, see [43, pages 8-10 and 8-11], or [10, corollary I.1.3.7].

Hence, if Γ is Kleinian, the quotient H^n/Γ is a

- complete
- convex
- hyperbolic
- 3-manifold
- with non-empty boundary.

Given M a convex manifold, define the **convex core** $H(M)$ to be the intersection of the submanifolds $N \subseteq M$ such that the canonical map $\pi_1(N) \to \pi_1(M)$ is an isomorphism. This is canonically identified with $(H\Lambda(\pi_1(M)))/\pi_1(M)$ (see also [38, p. 634]).

6.3.5 Proposition. *If M is a compact convex hyperbolic manifold, then there exists $\epsilon > 0$ such that any continuous deformation of M within an ϵ-neighbourhood of M can be enlarged in a small neighbourhood to give a convex hyperbolic manifold homeomorphic to M.*

Proof. Omitted, see [43, proposition 8.3.3], or [10, section I.2.5].

A manifold is **strictly convex** if every geodesic arc in M has interior a subset of $\operatorname{int} M$.

6.3.6 Proposition. *Let M_1 and M_2 be strictly convex manifolds, and let $\phi : M_1 \to M_2$ be a homotopy equivaence which is a diffeomorphism on ∂M_1. Then there exists a quasi-conformal homeomorphism $f : B^n \to B^n$ conjugating $\pi_1 M_1$ to $\pi_1 M_2$.*

Proof. Omitted, see [43, proposition 8.3.4].

6.3.7 Definition. Now let Γ be an arbitrary Kleinian group, acting naturally on H^n. We define

1. $M_\Gamma := (\text{h}-\text{conv}\,\Lambda(\Gamma))/\Gamma$ (the **convex hull quotient manifold**);
2. $N_\Gamma := H^n/\Gamma$ (the **complete hyperbolic manifold with boundary**);
3. $O_\Gamma := (H^n \cup \Omega(\Gamma))/\Gamma$ (the **Kleinian manifold**);

Let $W_\Gamma \subseteq \mathbb{P}^n$ be the set of points dual to planes in H^n which have intersection with S^{m-1}_∞ contained in $\Omega(\Gamma)$.

4 $P_\Gamma := (H^n \cup \Omega(\Gamma) \cup W_\Gamma)/\Gamma$ (the **completed Kleinian manifold**).

Observe that $H(N_\Gamma) = M_\Gamma \subseteq N_\Gamma \subseteq O_\Gamma \subseteq P_\Gamma$; that $M_\Gamma, N_\Gamma, O_\Gamma$ are homotopically equivalent; that $M_\Gamma \simeq_{\text{homeo.}} O_\Gamma$ except in degenerate cases; and that $N_\Gamma = \operatorname{int} O_\Gamma$.

6.4 The geometry of h−conv Λ(Γ)

We follow [43, section 8.5].

If $K \subseteq S^{n-1}_\infty$ is closed, then h−conv K is convex but each point on ∂ h−conv K lies on a hyperbolic geodesic segment in ∂ h−conv K. Thus, ∂ h−conv K develops to a hyperbolic plane. If Γ is torsion free then the hyperbolic structure projects well:

6.4.1 Proposition. *If Γ is a torsion free Kleinian group, then ∂M_Γ is a hyperbolic surface.*

Observe that ∂M_Γ is not flat (=is not a hyperbolic plane). Let $\gamma \subseteq \partial M_\Gamma$ be the set of points not in the interior of a flat region of M_Γ; we call γ the **bending locus**. Some properties:

- For all $x \in \gamma$, there exists some g_x a geodesic on ∂M_Γ through x (Fig. 6.7).
- γ is closed.
- If the area of ∂M_Γ is finite, then γ is compact.

This places a lamination structure on ∂M_Γ: the bending locus γ is the support of a geodesic lamination.

6.4.2 Theorem. *The set γ is not the entirety of $S := \partial M_\Gamma$ (in fact, it is measure 0).*

Proof. Observe that $S \setminus \gamma$ is a union of surfaces $\{S_i\}$ bounded by closed and/or infinite geodesics; each of these can be doubled along the boundary to a complete hyperbolic surface. The area of each S_i is bounded below by π — this follows from the Gauss-Bonnet theorem ([38, theorems 9.3.1 and 9.3.2]). Indeed, a complete manifold M (like the doubled surfaces) satisfies $\kappa\,\text{area}(M) = 2\pi\chi(M)$. In this case, $\kappa = -1$ and $\chi(M) < 0$, so $\text{area}(M) \geq 2\pi$. Since we doubled the S_i to obtain each M, we have $\text{area}(S_i) \geq \pi$ for all i.

Thus the number of components of $S \setminus \gamma$ is bounded above by $2|\chi(S \setminus \gamma)|$:

$$2\pi|\chi(S \setminus \gamma)| = \text{area}(S \setminus \gamma) = \sum_i \text{area}(S_i) \geq \pi(\text{number of components of } S \setminus \gamma);$$

We use the following lemma for a surface S with geodesic lamination supported on a set γ:

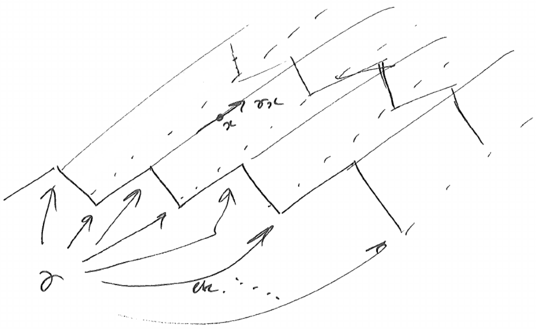


Figure 6.7: The pleated boundary of a convex hull quotient.

Lemma. $\chi(S) = \frac{1}{2}\chi(S - \gamma \textit{ doubled})$.

But again from Gauss-Bonnet, $\chi(S) = -\operatorname{area}(S)/2\pi$ and $\chi(S \setminus \gamma) = -\operatorname{area}(S \setminus \gamma \text{ doubled})/2\pi = -\operatorname{area}(S \setminus \gamma)/\pi$; thus substituting into the lemma we have $-\operatorname{area}(S)/2\pi = -\operatorname{area}(S \setminus \gamma)/2\pi$, so the measure of γ is 0.

Chapter 7

Geometrically finite groups

For this chapter, see [34, chapter VI], [38, chapter 12].

Recall that we defined 3-manifolds related to a Kleinian group in the previous chapter. We wish to restrict further to groups with 'finite tiling' actions; that is, each tile meets only finitely many others. More precisely, we wish the fundamental domains to be the interior of a shape made up of only finitely many hyperbolic line segments.

7.1 Fundamental polyhedra

7.1.1 Definition. Let G be Kleinian. An *open* polyhedron $D \subseteq H^3$ (i.e. the intersection of countably many hyperbolic open half-spaces $\{U_i\}_{i \in \mathbb{N}}$) is a **fundamental polyhedron** for D if the following criteria hold:

1. For every $K \subseteq H^3$ compact, only finitely many of the ∂U_i meet K;
2. For every $g \in G$ nontrivial, $g(D) \cap D = \emptyset$;
3. For every $x \in H^3$, there is a $g \in G$ with $g(x) \in \overline{D}$;
4. For every side s of D there is a side s' of D and an element $g_s \in G$ with $g_s(s) = s'$; further, $(s')' = s$ and $g_{s'} = g_s^{-1}$.
5. Any compact $K \subseteq H^3$ meets only finitely many G-translates of D.

In analogy with the dimension 2 case we shall use the words **facet** and **side** interchangeably for fundamental polyhedra.

Remark. The definition easily generalises for $n > 3$ to **fundamental polytopes**, and the following proposition also holds with minimal change in that case.

7.1.2 Proposition. *Let D be a fundamental polyhedron for a Kleinian group G. Then the interior B of $\overline{D} \cap \partial H^3$ (where $\overline{D}$ denotes the closure in $\overline{H^3}$) is a fundamental domain for G.*

Proof. We verify the conditions of Definition 4.2.1. In the following, diam A is the Euclidean diameter of some set A in the ball model. Observe that the proof of part 3 of the definition depends on part 5, so we prove them in that order below.

1. Suppose there are two elements of B equivalent under $g \in G$; then by continuity of the group action, $gD \cap D \neq \emptyset$, which is a contradiction unless $g = 1$. This shows that B is a G-packing.

2. If $z \in \Omega(G)$, then choose a sequence of points (x_n) in H^3 such that $x_n \to z$. For each n, there is some $g_n \in G$ such that $g_n x_n \in \overline{D}$; i.e. $x_n \in g_n^{-1}\overline{D}$. The sequence (g_m) must contain only finitely many distinct elements, otherwise diam $g_n^{-1}\overline{D} \to 0$ by Lemma 3.2.2 and so z is a limit point of the orbit of any point in $\overline{D}$, contradicting that $z \in \Omega$ via Theorem 3.3.27. Hence x_n is eventually z and $z \in g_n^{-1}(\overline{D})$ for some $g_n \in G$.

5. Let (s_n) be a sequence of sides of D; for each n, let P_n be the supporting hyperplane for s_n. Observe that diam $P_n \to 0$ since otherwise infinitely many of the P_n would meet some compact subset of H^3. Hence the diameter of the sides s_n goes to zero.

3 and 4. The sides of B are the boundaries of the facets of D, and the side-pairing transformations are the obvious ones. It remains to show that if $x \in B$ does not lie on a side then x is a limit point: suppose $x \in B$ does not lie on a side, then there is a sequence of facets (f_n) of D accumulating at x (it lies in the closure of a set of points lying on sides) and hence there is a sequence of side-pairing transformations (g_n) such that for all $z \in B$, $g_n(z) \to x$ (indeed, let g_n be the side pairing transformation *to* s_n for each n, where s_n is the side of B corresponding to g_n; since diam $s_n \to 0$, diam $g_n B \to 0$ and so $|g_n z - x| \to 0$).

6. Follows directly from (5) in the definition of a fundamental polyhedron.

Suppose $x_0 \in H^3$ is fixed by no nontrivial element of G. For every $g \in G$ nontrivial, the perpendicular bisector of the line joining x_0 to $g(x_0)$ is a hyperplane H_g in H^3. Let D_g be the half-space of points strictly closer to x_0 than to $g(x_0)$; then the **Dirichlet region** is the intersection of the D_g.

7.1.3 Theorem. *The Dirichlet region $D = \bigcap_{g \in G \setminus \{1\}} D_g$ is a fundamental polyhedron for G.*

Proof. That D is a countable intersection of halfspaces and is hence a polyhedron by our definition follows from Lemma 3.1.3. It remains to show the conditions of Definition 7.1.1.

1. Let $K \subseteq H^3$ be compact. There are only finitely many $g \in G$ such that $g(x_0) \in K$ (otherwise, let (g_n) be a sequence such that $g_n x_0 \in K$ for all n; then the sequence $g_n x_0$ accumulates at some $k \in K$ so k is a limit point of G, contradicting Theorem 3.3.27). Now note that if infinitely many of the H_g meet some compact subset L then a rotation and a dilation of L (both of which preserve compactness) will provide a compact set K containing all of the corresponding elements $g(x_0)$.

2. Let $g \in G$ be nontrivial. If $x \in D$, then $d(g(x), g(x_0)) = d(x, x_0) < d(x, g^{-1}(x_0)) = d(g(x), x_0)$ (the inequality coming from the definition of D, namely that x is closer to x_0 than to $g^{-1}(x_0)$) and so $g(x) \notin D$. Hence $gD \cap D = \emptyset$ and D is a G-packing.

3. Let $x \in H^3$; there is some $g \in G$ such that $d(x, g(x_0)) \leq d(x, h(x_0))$ for all $h \in G$ (that is, there is some translation of x_0 whose distance from x is minimal: otherwise, x would be a limit point). Given an arbitrary $h \in G$,

$$d(g^{-1}(x), x_0) = d(x, g(x_0)) \leq d(x, gh(x_0)) = d(g^{-1}(x), h(x_0));$$

in particular, $g^{-1}(x) \in \overline{D_h}$ for every $h \in G$ and so $g^{-1}(x) \in \overline{D}$.

4. Let $x \in \operatorname{relint} s$ for s a side of D. Then there is a unique $g \in G$ with $x \in \overline{D_g}$. This means that $d(x, x_0) < d(x, h(x_0))$ for all $h \neq g$, and $d(x, x_0) = d(x, g(x_0))$. Thus
$$d(g^{-1}(x), x_0) = d(x, g(x_0)) = d(x, x_0) = d(g^{-1}(x), g^{-1}(x_0))$$
and for any $h \neq g^{-1}$,
$$d(x, gh(x_0)) > d(x, x_0) = d(g^{-1}(x), x_0)$$
Hence $g^{-1}(x)$ lies on a side of D (say s'), and so $g^{-1}(s) = s'$. This sets up the side pairing transformations.

5. Finally, suppose $K \subseteq H^3$ is compact; if necessary expand K so that it is a closed ball around x_0, say of radius ρ. There are only finitely many translates of x_0 within K: if there were infinitely many that form a sequence $(g_n x_0)$, using compactness we can assume that the sequence converges to some $x \in K$ and so K contains a limit point, contradicting that it is a subset of H^3. If $d(g^{-1}(x_0), x_0) > 2\rho$, then $g(D) \cap K = \emptyset$. Thus the finitely many group elements which translate x_0 into K are the only group elements which map D to meet K. ⅅ

If G has a fundamental polyhedron in H^3 with only finitely many facets, then we say that G is **geometrically finite**.

7.2 Parabolic elements and punctures

A **horosphere** in B^n is a Euclidean sphere which is tangent to ∂B^n and which, apart from the point of tangency, lies within B^n. The horosphere is said to be **based** at the point of tangency. Such a sphere corresponds in H^n to either a Euclidean sphere in $\overline{H^n}$ tangent to ∂H^n, or (if it is based at ∞) to a Euclidean plane in H^n parallel to ∂H^n. The interior of a horosphere is called a **horoball**.

7.2.1 Lemma. *Let G be a Kleinian group containing the translation $z \mapsto z + 1$. Then the horoball*
$$T = \{z + tj \in H^3 \,:\, t > 1\}$$
is precisely invariant under $\operatorname{Stab}_G(\infty)$.

Proof. Let $J = \operatorname{Stab}_G(\infty)$. By Proposition 3.3.9, no element of J is loxodromic. In particular, every element of J is an Euclidean isometry, and the Poincaré extension of such a transformation of $\hat{\mathbb{C}}$ leaves horizontal planes invariant (in particular, the horoball of interest is a union of such planes and so is left invariant). Suppose that $g \in G$ is represented by the matrix
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix};$$
then by Proposition 3.4.1 either $c = 0$ so $g \in J$ or $|c| \geq 1$ and $g \notin J$; we prove that the latter elements move T off itself. Observe that the radius of the isometric circle $I(g)$ of g is $|c|^{-1} \leq 1$, and write $g = qr$ for q a Euclidean motion and r a reflection in $I(g)$. Taking the Poincaré extension, $\hat{r}$ is the reflection in a sphere of radius at most 1 about some point on $\mathbb{C}$; in particular, it moves T inside this sphere (and hence off itself); the following Euclidean motion does not change heights above $\mathbb{C}$ and so does not move T back onto itself. ⅅ

Recall that a **Fuchsian group** is a Kleinian group G such that G leaves some disc $D \subseteq \Omega$ invariant. Up to conjugation, we may assume that D is the upper half-plane H^2; in particular, G acts as a group of hyperbolic isometries on H^2.

7.2.2 Lemma. *If G is a Fuchsian group normalised to leave H^2 invariant and G does not fix ∞ then the isometric circle of G is centred at a point of $\mathbb{R}$.*

Proof. If G leaves H^2 invariant, then G leaves ∂H^2 invariant; in particular, preimages and images of ∞ lie in ∂H^2 and if they are not ∞ then they must be real.

7.2.3 Lemma. *Let G be a Fuchsian group containing the translation $z \mapsto z+1$. Then the horoball*

$$T = \{z \in H^2 : \operatorname{Im} z > 1\}$$

is precisely invariant under $\operatorname{Stab}_G(\infty)$.

Proof. Let $J = \operatorname{Stab}_G(\infty)$. By Proposition 3.3.9, no element of J is loxodromic. In particular, every element of J is a Euclidean isometry. Since G leaves H^2 invariant, these transformations must be compositions of translations $z \mapsto z + \alpha$ for $\alpha \in \mathbb{R}$ (which preserve T), or reflections along vertical lines (which preserve T). Thus $JT \subseteq T$.

Suppose that $g \in G$ is represented by the matrix

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

then by Proposition 3.4.1 either $c = 0$ so $g \in J$ or $|c| \geq 1$ and $g \notin J$; we prove that the latter elements move T off itself. Observe that the radius of the isometric circle $I(g)$ of g is $|c|^{-1} \leq 1$, and write $g = qr$ for q a Euclidean motion and r a reflection in $I(g)$. Since r is the reflection in a sphere of radius at most 1 about some point on $\mathbb{R}$ (Lemma 7.2.2), it moves T inside this sphere (and hence off itself); the following Euclidean motion does not change heights above $\mathbb{R}$ and so does not move T back onto itself.

7.2.4 Theorem. *Let G be a Fuchsian group acting on H^2 and containing the parabolic element $j : z \mapsto z+1$ such that if $g \in G$ satisfies $g^m = j$ for some $m > 0$, we have $g = j$ and $m = 1$. Then there is a punctured disc, conformally embedded in H^2/G, so that under the natural homomorphism $\pi_1((H^2 \cap {}^\circ\Omega)/G) \to G$ the element j corresponds to a loop about the puncture.*

Proof. Let $T = \{z \in \mathbb{C} : \operatorname{Im} z > 1\}$; since G is Fuchsian, $\operatorname{Stab}_G(T) = \langle j \rangle$ by Lemma 7.2.3. The map $f : \mathbb{C} \to \mathbb{C}$ given by $f(z) = \exp(2\pi i z)$ is a conformal map which sends T onto a punctured disc centred at 0; further it is a covering map with defining group $\langle j \rangle$, i.e. for $z, w \in T$, $f(z) = f(w)$ iff z and w are $\langle j \rangle$-equivalent.

Our goal is now the generalisation of Theorem 7.2.4 to the Kleinian case.

Let G be a Kleinian group, and let $J \leq G$ be an elementary group of parabolic type generated by a single element; let x be the unique fixed point of J. We say that J is **cusped** and that x is a **cusp point** of G if there is an open disc $B \subseteq \hat{\mathbb{C}}$ such that $J = \operatorname{Stab}_G(B)$ and such that B is precisely invariant under B. We say x is the **centre** of B, and that B is a **cusped region** for J.

7.2.5 Example. Let $G = J = \langle j = z \mapsto z/(z+1) \rangle$. The element j is parabolic (it has trace 2) with fixed point 0. What might the invariant discs be? Well, the element j normalises to $z \mapsto 1 + z$ upon conjugation by $z \mapsto 1/z$; an invariant disc for $z \mapsto 1 + z$ is H^2; and $1/(H^2)$ is $-H^2$.

Remark. Here is a geometric dictionary:

cusps	$\leftrightarrow$	punctures
cone points	$\leftrightarrow$	branch points (of finite order).

See [24] for an explanation of the name 'cone point'.

Similarly, we say that J is **doubly cusped** if there are two disjoint open discs $B_1, B_2 \subseteq \hat{\mathbb{C}}$ such that $B_1 \cup B_2$ is precisely invariant under J; we also say, in this case, that $B = B_1 \cup B_2$ is a **cusped region** for J. In the double cusp case it is clear that $\overline{B_1}$ and $\overline{B_2}$ are tangent at x: if B is precisely invariant under J then its closure must contain x otherwise we would not be able to find x as a limit point of the orbits of elements of B; a little further thought shows that this argument in fact implies that $x \in \overline{B_1} \cap \overline{B_2}$!

Remark. Compare with [35, p. 49].

7.2.6 Example. Let $J = G = \langle z \mapsto z + 1, z \mapsto -z \rangle$. Set $B_1 = \{z \in \mathbb{C} : \operatorname{Im} z > 1\}$ and $B_2 = \{z \in \mathbb{C} : \operatorname{Im} z < -1\}$. Clearly $B_1 \cup B_2$ is precisely invariant under J while each individual B_i is not.

Observe here that the unique fixed point of J is ∞, and the quotient $\Omega/G = \mathbb{C}/G$ is a sphere with two punctures and two cone points of angle π.

7.2.7 Example. Suppose G is a Fuchsian group with x a parabolic fixed point; then x is a double cusp point of G. Indeed, we may conjugate G such that the fixed point is ∞ and the fixed disc is H^2, and then use Theorem 7.2.4 to see that x is a double cusp point with two regions H^2 and $-H^2$.

Suppose that B is a cusped region for x; then there is some $h \in \mathbb{M}$ with $h(x) = \infty$ conjugating the generator of J to $z \mapsto z + 1$. The set $E = \{z : |\operatorname{Re} z| < 1/2\}$ is a fundamental domain for hJh^{-1}; then h^{-1} is a fundamental domain for J, and $h^{-1}(E) \cap B$ is called a **cusp** for J.

7.2.8 Proposition. *Let G be a geometrically finite Kleinian group with fundamental polyhedron D; let x be a point of $\overline{D} \cap \partial H^3$. Then either $x \in \Omega(G)$, or $J \coloneqq \operatorname{Stab}_G(x)$ is elementary of parabolic type. If J is cyclic, then x is doubly cusped.*

Proof.

Chapter 8

Moduli spaces of Kleinian groups

We follow closely the paper [28] to prove various continuity results about one-parameter holomorphic families of Kleinian groups: in particular, that the geodesic length function and the bending measure vary continuously with the parameter (exact statements may be found in Section 8.3).

This is the first of three chapters on the **Riley slice** and related moduli spaces.

8.1 Preliminaries on Kleinian group spaces

Let $D \subseteq \mathbb{C}$ be a connected domain; for each $\mu \in D$, let G_μ be a finitely generated Kleinian group *of the second kind* (so $\Omega(H_\mu) \neq \hat{\mathbb{C}}$). The system (G_μ) is a **holomorphic family of Kleinian groups** if we may choose a basepoint $\mu_0 \in D$ and a map $i : D \times \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ such that

- $i(\cdot, z)$ is holomorphic for all $z \in \hat{\mathbb{C}}$;
- $i(\mu, \cdot)$ is a quasi-conformal homeomorphism for each $\mu \in D$, such that the induced map

$$\begin{aligned} \phi_\mu : G_{\mu_0} &\to \mathbb{M} \\ g &\mapsto i_\mu g i_\mu^{-1} \end{aligned}$$

  is a type-preserving isomorphism onto G_μ.

Throughout, we shall be working with fixed connected components of $\Omega(G_\mu)$. For each $\mu \in D$, fix a connected component $\Omega^*(G_\mu)$; we will usually write $\Omega(\mu)$ for $\Omega(G_\mu)$, and $\Omega^*(\mu)$ for $\Omega^*(G_\mu)$.

Snce G_μ is finitely generated, by Ahlfors' finiteness theorem (Theorem 4.1.5) the space $\Omega^*(\mu)/\operatorname{Stab}_{G_\mu} \Omega^*(\mu)$ is a compact Riemann surface of finite genus with finitely many punctures.

Let $\mathcal{C}(\mu)$ for each μ denote h−conv $\Lambda(G_\mu)$, and let $\partial\mathcal{C}(\mu)$ be the boundary of $\mathcal{C}(\mu)$ in H^3. Each connnected component of $\partial\mathcal{C}(\mu)$ may be naturally identified with a component of $\Omega(\mu)$: the restriction of the retraction $r : \overline{H^n} \to \mathcal{C}(\mu)$ to each component of $\Omega(\mu)$ is a bijection onto a connected component of $\partial\mathcal{C}(\mu)$, and different components of $\Omega(\mu)$ are associated to different components of $\partial\mathcal{C}(\mu)$. We shall write $\partial\mathcal{C}^*(\mu)$ for the connected component corresponding to $\Omega^*(\mu)$.

We write β_μ for the induced transverse measure on $\partial\mathcal{C}^*(\mu)/\operatorname{Stab}\Omega^*(\mu)$ obtained by restricting and projecting the bending measure on $\partial\mathcal{C}(\mu)$. Let ℓ_μ denote the intrinsic measure on $\partial\mathcal{C}(G_\mu)$.

The following result is proved as [28, proposition 3.1] and relies on a result of Moore [37, 45]

8.1.1 Proposition. *The surfaces*

$$S_\mu := \partial\mathcal{C}(\mu)/G_\mu$$

and

$$\Omega^*(\mu)/\operatorname{Stab}\Omega^*(\mu)$$

are homeomorphic for all $\mu \in D$.

8.1.2 Corollary. *If $\mu, \nu \in D$, then S_μ and S_ν are quasiconformally homeomorphic.*

8.1.3 Corollary. *The map $i(\mu, \cdot)$ inducing the isomorphism $\phi_\mu : G_{\mu_0} \to G_\mu$ induces a homeomorphism $S_{\mu_0} \to S_\mu$ and an isomorphism $\pi_1(S_{\mu_0}) \to \pi_1(S_\mu)$.*

The inclusion $\Omega^*(\mu) \hookrightarrow H^3 \cup \Omega^*(\mu)$ induces a homomorphism

$$j_\mu : \pi_1\left(\frac{\Omega^*(\mu)}{\operatorname{Stab}\Omega^*(\mu)}\right) \simeq \pi_1(S_\mu) \to \pi_1\left(\frac{H^3 \cup \Omega^*(\mu)}{G_\mu}\right) \simeq G_\mu.$$

8.1.4 Proposition (Existence of conjugacy lift)**.** *There is a map*

$$R : D \times Z(\mu_0) \to Z(\mu)$$

(where $Z(\mu) := Z(G_\mu)$, the supporting hyperplane parameter space) such that for each $\mu \in D$, the induced map $R(\mu, \cdot)$ is a homeomorphism $Z(\mu_0) \to Z(\mu)$ with the additional property that the lifted map

$$R_{\mu *} : \pi_1(S_{\mu_0}) \to \pi_1(S_\mu)$$

makes the following diagram commute:

$$\begin{array}{ccc} \pi_1(S_{\mu_0}) & \xrightarrow{R_{\mu *}} & \pi_1(S_\mu) \\ \downarrow{\scriptstyle j_{\mu_0}} & & \downarrow{\scriptstyle j_\mu} \\ G_{\mu_0} & \xrightarrow{\phi_\mu} & G_\mu. \end{array}$$

Proof. Let $r_\mu : \Omega^*(\mu) \to \partial\mathcal{C}(\mu)$ be the retraction map; this lifts to a map $\hat{r}_\mu : \Omega^*(\mu) \to Z_\mu$ by sending $z \mapsto (r_\mu(z), P_r(z))$ where $P_r(z)$ is the support plane tangent to the horosphere based at z passing through $r_\mu(z)$. This lift is a homeomorphism. Set $R(\mu, P) := \hat{r}_\mu(i(\mu, \hat{r}_{\mu_0}^{-1}(P)))$ for all $P \in Z(\mu_0)$. This is a homeomorphism which conjugates the action of G_{μ_0} on Z_{μ_0} to the action of G_μ on Z_μ; a simple diagram chase shows the commutativity required.

8.2 Measured laminations

Let L be a geodesic lamination on a complete oriented hyperbolic surface S of finite area, and let ν be a transverse measure on L such that ν is preserved by isotopies moving one transversal to another which preserve the leaves of the lamination. We call the system (L, ν) a **measured lamination**.

If γ is a simple closed geodesic on S, then δ_γ denotes the measured lamination whose single leaf is the geodesic γ, and whose measure is an atomic unit mass on intervals transverse to γ.

Denote by $\mathcal{ML}(S)$ the space of all measured laminations on S. There is a natural topology on $\mathcal{ML}(S)$: if Σ is the set of free homotopy classes of simple closed curves on S, then we may define an embedding of $\mathcal{ML}(S)$ into $(\mathbb{R}_{>0})^\Sigma$ as follows: if $\nu \in \mathcal{ML}(S)$ and $[\sigma] \in \Sigma$ then define

$$\hat{\nu}([\sigma]) := \inf_{\sigma \in [\sigma]} \nu(\sigma)$$

where the infimum is taken over curves which are piecewise made up of transversal geodesics to ν or geodesics running along leaves of ν (such geodesics have measure zero by convention). Then the map

$$\nu \mapsto \hat{\nu}$$

is the required embedding. (This agrees with the weak topology on the measure space.)

If S has cusps, we use $\mathcal{ML}_0(S)$ to denote the set of measured laminations whose leaves end at cusps of S with the induced topology.

Restrict again to the situation where we have a moduli space of Kleinian groups; with the notation of the previous section,

8.2.1 Lemma. *For $\mu, \nu \in D$, the spaces $\mathcal{ML}(S_\mu)$ and $\mathcal{ML}(S_\nu)$ are homeomorphic.*

Proof. It suffices to show that $\mathcal{ML}(S_\mu)$ is homeomorphic to $\mathcal{ML}(S_{\mu_0})$ for μ_0 some fixed basepoint as above. If Γ is a family of pairwise disjoint geodesics on S_{μ_0} then R_μ maps this family to a family of quasi-geodesics (why?) and so we may associate to each geodesic γ_0 in Γ a geodesic in S_μ homotopic to $R_\mu(\gamma)$ in an invertible manner (since R_μ is a homeomorphism).

We write $\mathcal{ML}(S)$ for $\mathcal{ML}(S_{\mu_0})$ and identify this space, for each μ, with $\mathcal{ML}(S_\mu)$ under the homeomorphism defined in the lemma.

8.3 Statements of theorems on continuity

We now state the main theorems of [28]. Again we have a holomorphic family of Kleinian groups $\{G_\mu\}_{\mu\in D}$ and the associated notation. In addition, use ℓ_μ to denote the geodesic length measure with respect to the intrinsic metric on S_μ.

8.3.1 Theorem (Continuity of geodesic length)**.** *Recall that the S_μ are all homeomorphic, so $\pi_1(S_\mu)$ is independent of μ. For each $[\sigma] \in \pi_1(S_\mu)$, the length function*

$$\begin{array}{ccc} D & \longrightarrow & R_{\geq 0} \\ \mu & \longmapsto & \ell_\mu([\sigma]) \end{array}$$

is continuous.

8.3.2 Corollary. *The hyperbolic structure of S_μ varies continuously with μ.*

Proof. The Teichmüller space of a hyperbolic surface is embedded in $(\mathbb{R}_{>0})^\Sigma$ by the map

$$\mu \mapsto \{\ell_\mu([\sigma])\}_{[\sigma\in\Sigma]}$$

(where Σ is the set of free homotopy classes of simple closed curves on S_μ, again this is independent of μ).

8.3.3 Theorem (Continuity of bending measure)**.** *The map*

$$\begin{array}{ccc} D & \longrightarrow & \mathcal{ML}(S) \\ \mu & \longmapsto & \beta_\mu \end{array}$$

is continuous.

If $\nu \in \mathcal{ML}(S_\mu)$ then write $\ell_\mu(\nu)$ for the total mass of the measure on S_μ which is locally defined as the product of the measure ν on transversals to the lamination of ν and hyperbolic distance along the leaves of ν.

8.3.4 Theorem (Continuity of lamination length)**.** *The map*

$$\begin{aligned} D \times \mathcal{ML}_0(S) &\longrightarrow \mathbb{R}_{\geq 0} \\ (\mu, \nu) &\longmapsto \ell_\mu(\nu) \end{aligned}$$

is continuous in each variable.

8.4 Proofs of the continuity of geodesic length and bending measure

The goal of this section is the exposition of the proofs of Theorem 8.3.1 and Theorem 8.3.3. In this section, the propositions (which are named) are the main results which we shall use to prove the theorems; the lemmata are steps in the proofs of the propositions.

Notation. Let $\hat{\beta}_\mu$ and $\hat{\ell}_\mu$ denote the lifts of β and ℓ to $Z(\mu)$. We write:

1. $\rho_{\boxplus}(x, y)$ for the Euclidean distance between x and y in the halfspace model of H^3;
2. $\rho_\circ(x, y)$ for the Euclidean distance between x and y in the ball model of H^3 and area(Δ) for the Euclidean area of some shape Δ in the disc model of H^2;
3. $d(x, y)$ for the intrinsic hyperbolic metric in H^3 and h−area(Δ) for the hyperbolic area of some shape Δ in H^2;
4. $d_\omega(x, y)$ for the distance between x and y measured along some curve ω with respect to the intrinsic hyperbolic metric; more generally, d_X for the distance in some metric space X when needed.

Fix μ; we study finite approximations to the bending measure and geodesic lengths with respect to G_μ. Since continuity is a local property, we may restrict ourselves to the situation that μ varies in some compact $K \subseteq D$.

A polygonal approximation $(x_i, P_i)_{i=0}^n$ to a path ω in Z is an (α, S)-**approximation** if

$$\max_{1 \leq i \leq n} \theta(P_{i-1}, P_i) < \alpha \text{ and } \max_{1 \leq i \leq n} d_{\partial\mathcal{C}}(x_{i-1}, x_i) < s.$$

8.4.1 Proposition (Local error estimate)**.** *There is a universal constant K (independent of μ) and a function $s : [0, 2\pi) \to (0, 1)$ such that if $(x_i, P_i)_{i=0}^n$ is an $(\alpha, s(\alpha))$-approximation to a path ω in Z where $\alpha < \pi/2$, then*

1. $\left| \sum_{i=1}^n d_i(x_{i-1}, x_i) - \ell(\omega) \right| < K\alpha\ell(\omega)$*, and*
2. $\left| \sum_{i=1}^n \theta(P_{i-1}, P_i) - \beta(\omega) \right| < K\alpha\ell(\omega)$

where d_i denotes the induced metric on $P_{i-1} \cup P_i$ for each i.

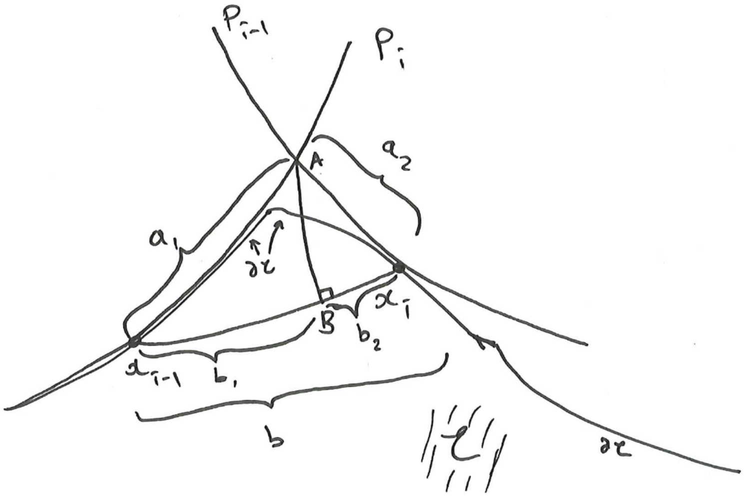


Figure 8.1: Computing, locally, the error of the geodesic length when taken along a polygonal approximation.

Proof. Fix a polygonal approximation $(x_i, P_i)_{i=0}^n$ for ω.

1. We consider the error $d_i - \ell_i$, where $d_i(x_{i-1}, x_i)$ and ℓ_i denotes the distance from x_{i-1} to x_i along ω. Choose a hyperbolic plane H through x_{i-1} and x_i such tht the shortest path between x_{i-1} and x_i in $P_{i-1} \cup P_i$ is contained in H; let a_1 and a_2 be the lengths of this shortest path in P_{i-1} and P_i respectively, and let b be the length of the geodesic $[x_{i-1}, x_i]$ in H^3. Let B be the orthogonal projection of the point $P_{i-1} \cap P_i \cap H$ onto γ, and let $b_1 = d(x_{i-1}, B)$ and $b_2 = d(B, x_i)$ such that $b = b_1 + b_2$. This notation is summarised in Fig. 8.1.

   Suppose the angle of the triangle $Ax_{i-1}x_i$ at x_{i-1} is $\varkappa_1$, and the angle at x_i is $\varkappa_2$. Then by hyperbolic trigonometry (in particular, [6, theorem 7.11.2]) we have $\tanh a_i = \tanh b_i \sec \varkappa_i$ for $i \in \{1, 2\}$.

   Since $\alpha < \pi/2$, and the acute angles of the triangle $Ax_{i-1}x_i$ are less than α (indeed, if these angles are $\varkappa_1, \varkappa_2$ then we have $(\pi - \alpha) + \varkappa + \lambda < \pi$, so $\varkappa + \lambda < \alpha$), we have that $\sec \varkappa_i < \sec \alpha$; thus

$$\tanh a_i < \tanh b_i \sec \alpha, \text{and} b_i < a_i < b \leq d_\omega(x_{i-1}, x_i).$$

   Now use a CAT(0)-space argument to see that there exists some choice of $s = s(\alpha)$ such that

$$d_\omega(x_{i-1}, x_i) < s \implies \tanh a_i > (1 - \alpha) a_i$$

   (namely, make s sufficiently small that the triangle $Ax_{i-1}x_i$ must have large side length to have the same fixed α). Using the standard inequality $\tanh b_i \leq b_i$ and applying the above results, we have

$$a_1 + a_2 < (b_1 + b_2) \sec \alpha + \alpha(a_1 + a_2)$$

   and hence, since $\sec \alpha \geq 1$,

$$a_1 + a_2 - b < K\alpha d_\omega(x_{i-1}, x_i).$$

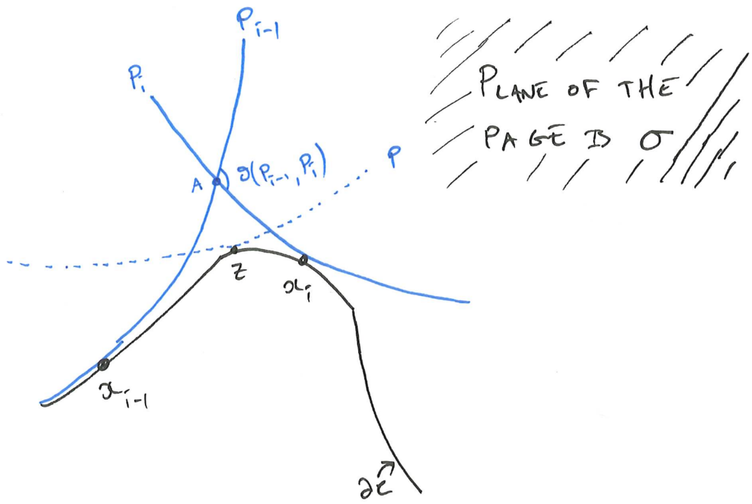


Figure 8.2: Computing, locally, the error of the bending measure when taken along a polygonal approximation.

2. Let $t \in [0,1]$ be an intermediate partition point, so $t_{i-1} < t < t_i$; let $(z, P) = \omega(t)$. Apply Lemma 5.4.1 to the triplet of planes P_{i-1}, P, P_i: there is either a unique horosphere σ through x_{i-1} and orthogonal to each of the three planes, or there is a unique hyperbolic plane σ orthogonal to each of them. Let $l_{i-1} = \sigma \cap P_{i-1}$, $l = \sigma \cap P$, and $l_i = \sigma \cap P_i$. For the notation, see Fig. 8.2.

   Observe that the angle sum of the approximation before refinement is

$$\theta(P_0, P_1) + \cdots + \theta(P_{i-2}, P_{i-1}) + \theta(P_{i-1}, P_i) + \theta(P_i, P_{i+1}) + \cdots + \theta(P_{n-1}, P_n)$$

   and after refinement it is

$$\theta(P_0, P_1) + \cdots + \theta(P_{i-2}, P_{i-1}) + \theta(P_{i-1}, P) + \theta(P, P_i) + \theta(P_i, P_{i+1}) + \cdots + \theta(P_{n-1}, P_n);$$

   since refinement decreases the approximation, the change in the approximation is

$$\epsilon = \theta(P_{i-1}, P_i) - \theta(P_{i-1}, P) - \theta(P, P_i).$$

   We have two cases:

   - If σ is a horosphere (so has Euclidean intrinsic geometry), then the triangle formed by l_{i-1}, l, and l_i is Euclidean and thus the error ε is zero (consider Fig. 8.3).
   - If σ is a hyperbolic plane, then the error is

$$\theta(P_{i-1}, P_i) - \theta(P_{i-1}, P) - \theta(P, P_i) = \pi - (\pi - \theta(P_{i-1}, P_i)) - \theta(P_{i-1}, P) - \theta(P, P_i);$$

     that is, ϵ is π minus the sum of the triangle angles; it is a standard result in hyperbolic geometry (see for instance [6, theorem 7.13.1]) that this is exactly the hyperbolic area of the triangle.

Using this argument, we shall prove the following intermediate claim:

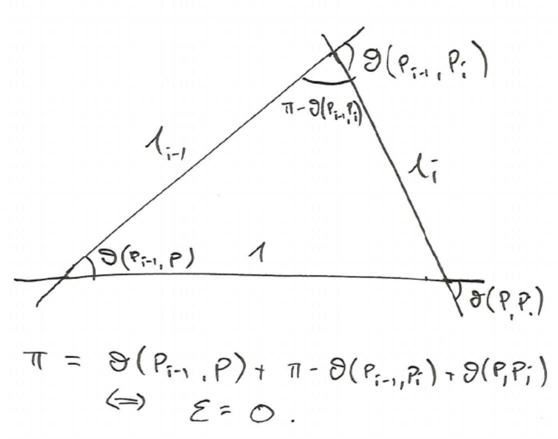


Figure 8.3: Change in the error of an approximation under a refinement.

Claim. *The hyperbolic area* h−area(R) *is an upper bound for the error* $\theta(P_{i-1},P_i) - \beta(\omega_i)$ *where* ω_i *is the portion of* ω *between* t_{i-1} *and* t_i*, and where* R *is the region bounded by the curves* l_i*,* l_{i-1}*, and* ω_i*.*

Indeed, recall that $\beta(\omega_i)$ is the infimum over all the permissible refinements $t_{i-1} = s_0 < \cdots < s_m = t_i$ of $[t_{i-1}, t_i]$ of the quantity $\sum_{j=1}^{m} \theta(Q_{j-1}, Q_j)$ where Q_j denotes the second component of $\omega(s_j)$. It is therefore enough to show that h−area(R) is an upper bound for the quantity $\theta(P_{i-1}, P_i) - \sum_{j=1}^{m} \theta(Q_{j-1}, Q_j)$ for every such refinement; and since these refinements are finite by definition, we may prove this claim by induction. The base case was done above: denote by Δ_0 the triangle with edges l_{i-1}, l, and l_i. Assume now that the claim is true for refinements of $m-1$ points, and consider the partition $t_{i-1} = s_0 < \cdots < s_m = t_i$ as above. Removing the point x_1, we have by induction a disjoint union of triangles $\Delta_0, ..., \Delta_m$ whose area is an upper bound (by doing the argument above for each triangle) for the error in the approximation given the partition of $m-1$ points. Adding back in x_1 and forming the plane σ_m as above (using Lemma 5.4.1), the change in the approximation is bounded above (by the argument given above) by the area of the triangle formed by $Q_1 \cap \sigma_m$, $Q_0 \cap \sigma_m$, and $Q_2 \cap \sigma_m$, which is in turn bounded above by the area of the triangle Δ_m bounded by $Q_1 \cap \sigma$, $Q_0 \cap \sigma$, and $Q_2 \cap \sigma$ (since the former is an orthogonal projection of the latter). Hence the error $\theta(P_{i-1}, P_i) - \sum_{j=1}^{m} \theta(Q_{j-1}, Q_j)$ increases by at most the area of Δ_m; i.e.

$$\theta(P_{i-1}, P_i) - \sum_{j=1}^{m} \theta(Q_{j-1}, Q_j) \leq \text{h−area}\,\Delta_0 + \cdots + \text{h−area}\,\Delta_m \leq \text{h−area}(R)$$

as desired. This proves the intermediate claim.

We now prove a bound on h−area(R) by comparing it with the Euclidean area area(R) in the disc model. Map the plane σ into the unit disc using the standard identification of H^2 with B^2, sending $A \mapsto 0$; call this map ϕ. By convexity, the region $\phi(R)$ is contained within the (Euclidean) triangle Δ' with vertices 0, $\phi(x_{i-1})$, and ϕ_i. By assumption, $d_\omega(x_{i-1}, x_i) < s$; since the hyperbolic distance d is a lower bound on the distance along ω, $d(x_{i-1}, x_i) < s$; and hence $d_\sigma(x'_{i-1}, x'_i) < s$, where x'_{i-1} and x'_i are the projections of x_{i-1} and x_i onto σ. Since $\alpha < \pi/2$,

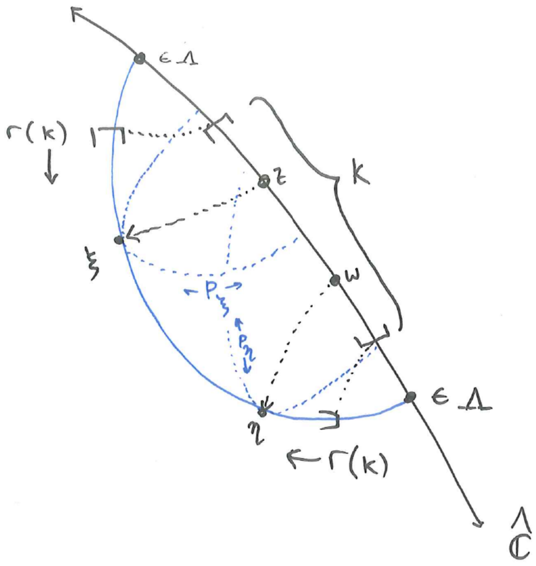


Figure 8.4: Images of z and w under $\hat{r}$ in the proof of Lemma 8.4.3.

the triangle Δ' is contained within a circle of bounded Euclidean radius in B^2; and thus there is a bounded comparison (say with constant L) between the Euclidean and hyperbolic areas of Δ'. Choose $s(\alpha)$ sufficiently small now such that the Euclidean distance $\rho_\circ(\phi(x_{i-1}), \phi(x_i))$ is at most 1; then the Euclidean triangle has area bounded above by $d(x_{i-1}, x_i)\alpha$. In particular,

$$\theta(P_{i-1}, P_i) - \beta(\omega_i) \le \text{area}(R) \le L\,\text{area}(\Delta') \le L\alpha\rho_\circ(\phi(x_{i-1}), \phi(x_i)) \le L\alpha d_\omega(\phi(x_{i-1}), \phi(x_i))$$

as required.

8.4.2 Proposition (Continuity of conjugacy lift). *If $\{G_\mu\}_{\mu\in D}$ is a holomorphic family of Kleinian groups, then the map*

$$R_\cdot : Z_{\mu_0} \times D \to Z_\mu$$

defined (as in Proposition 8.1.4) by

$$(\omega, \mu) \mapsto (\hat{r}_\mu i_\mu \hat{r}_{\mu_0}^{-1})(\omega)$$

is continuous when the domain is equipped with the usual product topology.

The proof of Proposition 8.4.2 will depend on the following lemmata.

8.4.3 Lemma. *Let $\Lambda \subseteq \mathbb{C}$ be an arbitrary closed set with diameter greater than some constant $c > 0$; let K be a closed bounded convex subset of a connected component of $\mathbb{C} \setminus \Lambda$. Let r be the retraction map $\overline{H^3} \to$ h−conv Λ. Then $\hat{r}$ is uniformly continuous on K, where the modulus of continuity*[1] *depends only on $a = d_{\mathbb{H}}(K, \Lambda)$, $b = \sup_{z\in K} d_{\mathbb{H}}(z, \Lambda)$, and c.*

Proof. Let $z, w \in K$ and let $\hat{r}(z) = (\xi, P_\xi)$ and $\hat{r}(w) = (\eta, P_\eta)$. Since z and w lie in the same component of $\hat{\mathbb{C}} \setminus \Lambda$, the hemispheres bounded by P_ξ and P_η cannot be nested (Fig. 8.4).

[1] Recall that the **modulus of continuity** for a function on metric spaces $f : (X_1, d_1) \to (X_2, d_2)$ is an increasing real-valued function $\omega : \hat{\mathbb{R}}_{\ge 0} \to \hat{\mathbb{R}}_{\ge 0}$ vanishing and continuous at 0 such that for all $x, y \in X_1$, $d_2(f(x), f(y)) \le \omega(d_1(x, y))$.

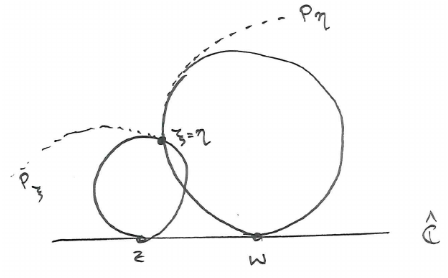


Figure 8.5: The case $\xi = \eta$ but $P_\xi \neq P_\eta$ may occur if $\xi = \eta$ lies on a pleating edge.

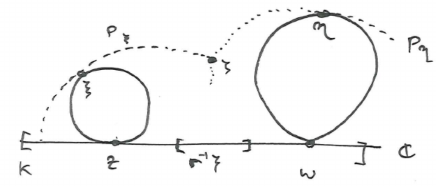


Figure 8.6: The case $\xi = \eta$ and $P_\xi \neq P_\eta$ may occur if ξ and η lie on adjoining flat pieces.

Since K is bounded, there exist positive constants a' and b' depending only on a, b, and c such that for all choices of $z, w \in K$,

$$a' \leq \text{height}\,\xi \leq b' \qquad a' \leq \text{height}\,\eta \leq b'$$
$$a' \leq \text{diam}\,S(\xi) \leq b' \quad a' \leq \text{diam}\,S(\eta) \leq b'$$

where height is the height function in the half-space model, and where $S(\xi)$ and $S(\eta)$ are the horospheres tangent to P_ξ and P_η respectively.

In particular we have four bounds:

1. If $P_\xi = P_\eta$ then $d_{\mathbb{H}}(\xi, \eta)$ depends only on z and w, and tends uniformly to 0 as $|z - w| \to 0$ with constants depending only on a' and b'. Indeed, if $\pi\xi$ and $\pi\eta$ are the orthogonal projections of ξ and η onto $\mathbb{C}$, then
$$d_{\mathbb{H}}(\pi\xi, \pi\eta) \leq d_{\mathbb{H}}(z, w)$$
(since π is distance-reducing) and so $d_{\mathbb{H}}(\pi\xi, \pi\eta)$ depends uniformly only on $d(z, w)$. Hence
$$[d_{\mathbb{H}}(\xi, \eta)]^2 = [\text{height}(\xi - \eta)]^2 + [d_{\mathbb{H}}(\pi\xi, \pi\eta)]^2$$
and height$(\xi - \eta)$ is also uniform in z and w with bounds depending on a' and b'.

2. If $\xi = \eta$ but $P_\xi \neq P_\eta$ (for instance this may occur if $\xi = \eta$ lies on a pleating edge, see Fig. 8.5)

   By similar arguments to (1), we see that $|z - w|$ tends to zero iff $d_Z((\xi, P_\xi), (\eta, P_\eta))$ tends to zero, uniformly with constants depending on a' and b'.

3. If $\xi \neq \eta$ and $P_\xi \neq P_\eta$ but $P_\xi \cap P_\eta \neq \emptyset$ (see Fig. 8.6), then by convexity of K there is a point $\zeta \in P_\xi \cap P_\eta$ such that $r^{-1}(\zeta) \subseteq K$.

   Choose some $u, v \in r^{-1}(\zeta)$ such that $\hat{r}(u) = (\zeta, P_\xi)$ and $\hat{r}(v) = (\zeta, P_\eta)$. Apply (1) and (2) above to the pairs (z, u) (case 1), (u, v) (case 2), and (v, w) (case 1) to see that $|z - w|$ tends to zero iff $d_Z((\xi, P_\xi), (\eta, P_\eta))$ tends to zero, uniformly with constants depending on a' and b'.

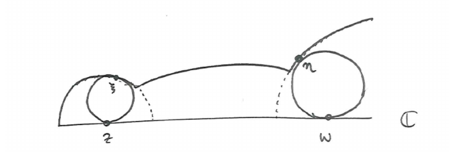


Figure 8.7: The case $P_\xi \cap P_\eta = \emptyset$ does not have any effect on convergence.

4. Finally, consider the case that $P_\xi \cap P_\eta = \emptyset$. In this case, $|z-w|$ is uniformly bounded below by height$(\xi - \eta)$ (which is in turn bounded below since ξ and η lie above the same connected component of Ω — see Fig. 8.7) and so we cannot send $|z-w| \to 0$.

Combining the cases 1–3 gives the desired convergence.

8.4.4 Lemma. *The function $\mu \mapsto d(K, \Lambda(\mu))$ is continuous at μ_0.*

Proof. By definition, $\Lambda(\mu) = i_\mu \Lambda(\mu_0)$ and so we wish to estimate

$$\left|d(K,\Lambda(\mu)) - d(K,\Lambda(\mu_0))\right| = \left|d(K, i_\mu\Lambda(\mu_0)) - d(K,\Lambda(\mu_0))\right|;$$

observe that $d(K, i_\mu \Lambda(\mu_0)) \le d(K,\Lambda(\mu_0)) + d(\Lambda(\mu_0), i_\mu(\Lambda(\mu_0)))$ and so $d(K,\Lambda(\mu)) - d(K,\Lambda(\mu_0)) \le d(\Lambda(\mu_0), i_\mu(\Lambda(\mu_0)))$; by uniform continuity of i_μ in μ in a small neighbourhood of μ_0, the right hand side can be bounded below ε for μ sufficiently close to μ_0. A similar estimate holds for the lower inequality.[2]

8.4.5 Lemma. *The function $\mu \mapsto \operatorname{diam}\Lambda(\mu)$ is continuous at μ_0.*

Proof. Let $\varepsilon > 0$. Let K be a compact neighbourhood of μ_0; observe that $i : K \times \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is uniformly continuous by the Heine-Cantor theorem. Hence we may pick $\delta > 0$ such that for all $\mu \in K$ and all $x \in \hat{\mathbb{C}}$, if $|\mu - \mu_0| < \delta$ then $|i(\mu,x) - i(\mu_0,x)| < \varepsilon$.

Now observe that

$$\begin{aligned}
\operatorname{diam}\Lambda(\mu) &= \sup_{x,y\in\Lambda(\mu_0)} \left|i(\mu,x) - i(\mu,y)\right| \\
&\le \sup_{x,y\in\Lambda(\mu_0)} \left(\left|i(\mu,x) - i(\mu_0,x)\right| + \left|-i(\mu,y) + i(\mu_0,y)\right| + \left|i(\mu_0,x) - i(\mu_0,y)\right|\right) \\
&\le 2 \sup_{x\in\Lambda(\mu_0)} \left|i(\mu,x) - i(\mu_0,x)\right| + \sup_{x,y\in\Lambda(\mu_0)} \left|i(\mu_0,x) - i(\mu_0,y)\right| \\
&= 2 \sup_{x\in\Lambda(\mu_0)} \left|i(\mu,x) - i(\mu_0,x)\right| + \operatorname{diam}\Lambda(\mu_0)
\end{aligned}$$

and so

$$\left|\operatorname{diam}\Lambda(\mu) - \operatorname{diam}\Lambda(\mu_0)\right| \le \operatorname{diam}\Lambda(\mu) - \operatorname{diam}\Lambda(\mu_0) \le 2\varepsilon$$

when μ is chosen δ-close to μ_0 as above.

8.4.6 Lemma. *Let $\{G_\mu\}_{\mu\in D}$ be a holomorphic family of Kleinian groups and pick a basepoint $\mu_0 \in D$. Without loss of generality, assume that the component $\Omega^*(\mu)$ is bounded in $\mathbb{C}$ (if $\infty \in \Omega^*(\mu)$ then replace G with its conjugate under a transformation moving ∞ into ΛG_μ). Let $K \subseteq \Omega^*(\mu_0)$ be convex and compact. Then, for every $\varepsilon > 0$, there exists $\delta > 0$ depending only on ε and $d(K, \Lambda\mu_0)$ such that*

[2] I should add the general principles of this to the chapter on hyperbolic convexity... in fact there should be a 'chapter 0' on hyperbolic geometry and convexity.

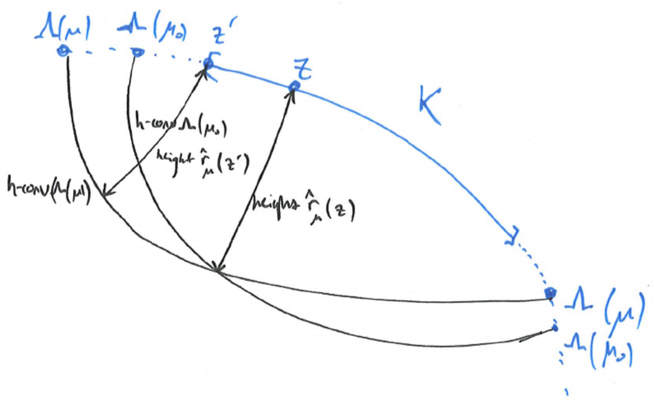


Figure 8.8: Heights above retractions of h−conv $\Lambda(\mu)$.

1. *$K \subseteq \Omega^*(\mu)$ whenever $|\mu - \mu_0| < \delta$; and*
2. *$d_Z(\hat{r}_\mu(z), \hat{r}_{\mu_0}(z)) < \varepsilon$ whenever $|\mu - \mu_0| < \delta$ and $z \in K$*

(where d_Z is the metric on $H^3 \times \mathbb{G}_2(H^3)$ which induces the Euclidean distance on the first component and the dihedral angle metric on the second).

Proof. By Lemma 8.4.4 and Lemma 8.4.5, we may pick some $\delta_0 > 0$ such that $d(K, \Lambda(\mu)) > d(K, \Lambda(\mu_0))/2$ and $\operatorname{diam}\Lambda(\mu) > \operatorname{diam}\Lambda(\mu_0)/2$ whenever $|\mu - \mu_0| < \delta_0$. Hence, since $d(K, w) \geq d(K, \Lambda(\mu))$ for all $w \notin \Omega^*(\mu)$, we have that $K \subset \Omega^*(\mu)$ for all such μ.

Assume now that $\operatorname{diam}\Lambda(\mu)$ is bounded below by $c > 0$, and let $z \in K$ be arbitrary. Since $d(K, \Lambda(\mu)) > d(K, \Lambda(\mu_0))/2$ for μ δ_0-close to μ_0, there exists (by a compactness argument) a constant a' independent of z such that height $r_\mu(z) \geq a'$ (Fig. 8.8)

Let $\hat{r}_0(z) = (\xi, P_0)$, and let γ be the geodesic extending $[\xi, z]$. Let H_t, for $t \in (-\text{height}\,\xi, \infty)$, be the horosphere based at z through the point on γ with signed distance t from ξ, measured such that the values of t corresponding to points on $[\xi, z]$ are negative. □

Proof of Proposition 8.4.2. □

Chapter 9

Combination theorems

The notes in this chapter are based on the lectures by Jeroen.

9.1 Amalgamated free products

Let G_1, G_2 be subgroups of some fixed universal group Γ, and let $J \leq \Gamma$ be a shared subgroup of both G_1 and G_2 such that $[G_m : J] > 1$ for $m \in \{1, 2\}$. We say that a word in the elements of $G_1 \cup G_2$ is a **normal form** if it is of the form $g_m \cdots g_1$ where either

- for all i odd, $g_i \in G_1 \setminus J$; and for all i even, $g_i \in G_2 \setminus J$; or
- for all i odd, $g_i \in G_2 \setminus J$; and for all i even, $g_i \in G_1 \setminus J$.

We place an equivalence relation on the space of normal forms; for all $j \in J$ and all $1 < k \leq n$, we say

$$g_n \cdots g_1 \sim g_n \cdots (g_k j)(j^{-1} g_{k-1}) \cdots g_1.$$

(Observe that $g_k j$ must lie in G_m if g_k lies in G_m, and cannot be an element of J as otherwise $g_k = (g_k j) j^{-1}$ is an element of J; similarly, $j^{-1} g_{k-1} \in G_{3-m}$ and so the word on the right is indeed a normal form.)

We say that the normal form $g_n \cdots g_1$ is a m-**form** if $g_n \in G_n$.

9.1.1 Lemma. *If the normal forms φ and ϑ are equivalent, and φ is a m-form, then ϑ is an m-form.*

We impose a multiplication on normal forms as follows: the product of two forms $\varphi = g_n \cdots g_1$ and $\vartheta = h_k \cdots h_1$ is the word obtained by juxtaposing $\varphi.\vartheta$; this is clearly a normal form, unless $g_1 h_k \in J$ in which case either $k = n = 1$ and $\varphi.\vartheta = g_1 h_k \in J$, or one of $g_2(h_1 h_k)$ or $(g_1 h_k) h_{k-1}$ lies in some $G_m \setminus J$ (being the same type as g_2 or h_{k-1} respectively if either of these exists). In particular, $\varphi.\vartheta$ is either an element of J or a normal form. The **amalgamated free product** of G_1 and G_2 along J is the group $G_1 *_J G_2$ supported on the union of the set of equivalence classes of words normal forms and the set of elements of J, with the natural extension of this multiplication.

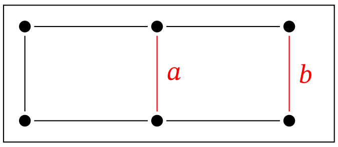
Remark. An alternative description of the amalgamated product is found in [41, section I.1].

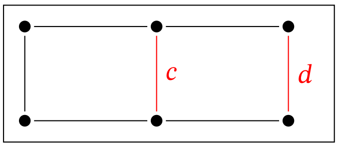
9.1.2 Example. $\mathbb{Z} * \mathbb{Z} \simeq F_2$ (here we adhere to the convention that if the common subgroup is omitted from the notation, it is implicitly understood to be 1).

9.1.3 Example. For this example we use the following theorem from algebraic topology.

Theorem ([33, theorem 10.12])**.** *Let X be a finite connected graph with maximal spanning tree T; then $\pi_1(X)$ is a free group generated by elements in bijective correspondence with the edges not in T.*

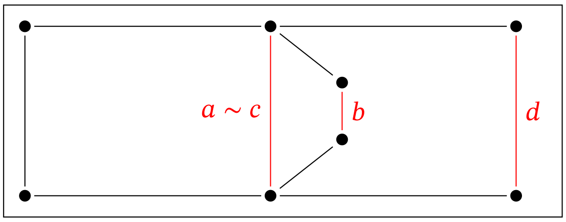
Let X and Y be the disjoint graphs

$X =$ 

; $Y =$ 

.

In each case, the graph induced by the black edges is a maximal spanning tree and so $\pi_1(X) \simeq \langle a, b\rangle$ and $\pi_1(Y) \simeq \langle c, d\rangle$. Consider the graph Z obtained by gluing X and Y along the 4-cycles containing a and c respectively:

$Z =$ 

.

Again the black edges induce a spanning tree, so the fundamental group of Z is the amalgamated product of $\pi_1(X)$ and $\pi_1(Y)$ along the subgroup J obtained by identifying $\langle a\rangle \le \pi_1(X)$ with $\langle c\rangle \le \pi_1(Y)$. More abstractly, $\pi_1(Z) = F_2 *_{\mathbb{Z}} F_2$.

This is a special case of the Seifert-Van Kampen theorem [33, theorem 10.1].

9.1.4 Example (Advanced example)**.** $\mathrm{SL}(2,\mathbb{Z}) \simeq \mathbb{Z}/4\mathbb{Z} *_{\mathbb{Z}/2\mathbb{Z}} \mathbb{Z}/6\mathbb{Z}$ [41, section I.4.2].

Let Φ denote the natural group homomorphism

$$\Phi : G_1 *_J G_2 \to G = \langle G_1, G_2\rangle$$
$$\Phi(g) = \begin{cases} g_n \cdots g_1 & \text{when } g \text{ is the normal form } g_n \cdots g_1; \\ g & \text{when } g \in J. \end{cases}$$

If $\vartheta = g_n \cdots g_1$ is a normal form, we set $|\vartheta| = n$ and call this number the **length** of the normal form.

9.1.5 Lemma. *If the normal forms φ and ϑ are equivalent, then $|\varphi| = |\vartheta|$.*

Observe that normal forms are not necessarily unique unless the amalgamated product is actually a free product (i.e. $J = 1$).

9.1.6 Lemma. *Suppose $G = G_1 *_J G_2$ and $\tilde{G} = \tilde{G_1} *_{\tilde{J}} \tilde{G_2}$. If $\phi_1 : G_1 \to G_2$ and $\phi_2 : \tilde{G_1} \to \tilde{G_2}$ are isomorphisms such that $\phi_1\restriction_J = \phi_2\restriction_J$ is an isomorphism $J \to \tilde{J}$, then there exists a unique $\phi : G \to \tilde{G}$ with $\phi\restriction_{G_1} = \phi_1$ and $\phi\restriction_{G_2} = \phi_2$.*

Our goal is a criterion for writing G as an amalgamated free product. We will actually do something less impressive, namely develop a criterion for writing a group $G = \langle G_1, G_2\rangle$ which acts freely discontinuously on X as an amalgamated free product.

We shall be proving a version of the **ping-pong lemma**, the philosophy of which is often attributed to Fricke and Klein.

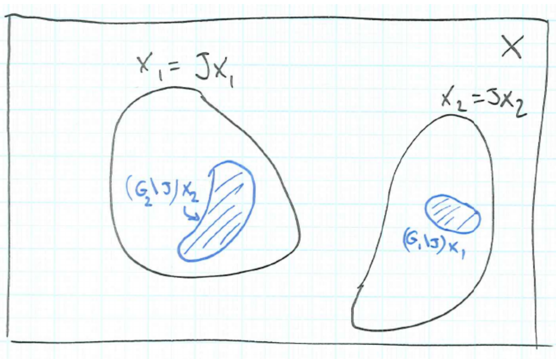


Figure 9.1: A generic interactive pair.

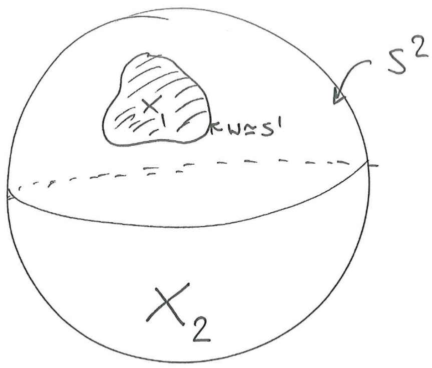


Figure 9.2: An interactive pair on S^3.

9.1.7 Definition. Suppose $G = \langle G_1, G_2\rangle$ and J is a shared proper subgroup of G_1 and G_2. Suppose G acts on a set X.

A pair (X_1, X_2) of subsets of X is an **interactive pair** if the following hold (compare Fig. 9.1):

1. $X_1 \neq \emptyset, X_2 \neq \emptyset$, and $X_1 \cap X_2 = \emptyset$;
2. $JX_1 \subseteq X_1$ and $JX_2 \subseteq X_2$; and
3. $(G_1 \setminus J)X_1 \subseteq X_2$ and $(G_2 \setminus J)X_2 \subseteq X_1$.

9.1.8 Proposition. *Let $X = S^n$, $G = \mathrm{M}(n)$ the group of Möbius transformations with the natural action on the unit sphere via stereographic projection (see chapter 1), and let W be a topological $(n-1)$-sphere on X bounding two open discs X_1 and X_2 (see Fig. 9.2). Suppose $G = \langle G_1, G_2\rangle$ with J a shared proper subgroup of both G_1 and G_2.*

If each X_m is precisely invariant under J in G_m, then (X_1, X_2) is an interactive pair.

Proof. Take $g \in G_1 \setminus J$. Since X_1 is J-invariant, $\overline{X_1}$ is J-invariant (by continuity). We have $gX_1 \cap X_1 = \emptyset$ for all such g and so since $W = \partial X_1$ we have $gW \cap X_1 = \emptyset$. Then $g(X_1) \subseteq X_2$ by inspection of Fig. 9.2; similarly $g(X_2) \subseteq X_1$.

If Γ is a group acting on a space X, by the axiom of choice there is a set $D \subseteq X$ containing exactly one representative of each orbit of Γ/X. We call such a set a **fundamental set**.

9.1.9 Theorem. *If G acts freely discontinuously on a non-empty open set $U \subseteq X$, and if $G = G_1 *_J G_2$, then there exists an interactive pair for the action.*

Proof. By the axiom of choice, there exists a fundamental set $D \subseteq U$ containing exactly one representative for each orbit of G in X. Let $X_1 = \bigcup g(D)$ where the union is taken over 2-forms and let $X_2 = \bigcup g(D)$ where the union is taken over 1-forms. Both X_1 and X_2 are non-empty since there exist 1- and 2-forms; both X_1 and X_2 are J-invariant, since if $g(d) \in g(D)$ for some $d \in D$ and some g an m-form, then jg is also an m-form and so $jg(d) = (jg)(d) \in X_m$.

We now show that $X_1 \cap X_2 = \emptyset$. Suppose for the sake of contradiction that $x \in X_1 \cap X_2$. Then $x = h_1(y_1) = h_2(y_2)$ for some $y_1, y_2 \in D$ and some $h_1, h_2 \in G$ respectively a 1- and a 2-form. Hence $y_1 = y_2 =: y$ since otherwise y_1, y_2 would be two distinct orbit representatives in D. Write $h_1 = g_n \cdots g_1$ and $h_2 = f_k \cdots f_1$; by construction $h_1^{-1}h_2(y) = y$ and so since $y \in {}^\circ\Omega$ we have $h_1^{-1}h_2 = 1$. On the other hand, $h_1^{-1}h_2 = g_1^{-1} \cdots g_n^{-1} f_k \cdots f_1$; since $g_h^{-1} \in G_1 \setminus J$ and $f_k \in G_2 \setminus J$, so $h_1^{-1}h_2$ is a normal form, thus not the identity (contradiction.) ⅅ

We remark that the converse is not true:

9.1.10 Example. Define subgroups of $\mathbb{M}$ as follows:

$$\begin{array}{cc} j(z) = z + 1 & J = \langle j \rangle \\ g_1(z) = -z & g_{2,a} = -z + 2a \ (a > 0) \\ G_1 = \langle j, g_1 \rangle & G_{2,a} = \langle j, g_{2,a} \rangle. \end{array}$$

Then $G_1 \simeq D_\infty$ and $G_{2,a}$ is a conjugate of G_1 in $\mathbb{M}$.

Now we show that there exists an interactive pair for the action of G with respect to the data $G_1 \geq J \leq G_2$, but that G is not the amalgamated product $G_1 *_J G_2$. Let $X_1 = H^2$ be the upper half-plane and $X_2 = H^{2-}$ the lower half-plane of $\mathbb{C}$.

Observe that X_1 is precisely invariant under J in G_1 and that X_2 is precisely invariant under J in $G_{2,a}$. Thus (X_1, X_2) is an interactive pair.

Suppose a is irrational. Then $G_{2,a}$ is not discrete, so the action is not freely discontinuous. Suppose a is rational. Then some power of $g_1 g_{2,a}$ lies in J, but $g_1 g_{2,a}$ is an alternating nontrivial product of elements not in J.

We say that $\vartheta = g_n \cdots g_1$ is a (m,k)-**form** if $g_n \in G_m \setminus J$ and $g_1 \in G_k \setminus J$. We say that an interactive pair (X_1, X_2) is **proper** if either

- there exists a point in X_1 not G_2-equivalent to any point in X_2, or
- there exists a point in X_2 not G_1-equivalent to any point in X_1.

The claim is that this condition is enough to ensure that the situation of Example 9.1.10 does not occur, and in fact we get a converse to Theorem 9.1.9.

We hide the difficulties in the following lemma.

9.1.11 Lemma. *Suppose $G = \langle G_1, G_2 \rangle$ acts on X and suppose $G_1 > J < G_2$ is a shared subgroup. If there exists an interactive pair (X_1, X_2) for this data, and if $\vartheta = g_n \cdots g_1$ is an (m,k)-form, then $\Phi(g)(X_k) \subseteq X_{3-m}$. Further, this inclusion is proper if (X_1, X_2) is proper and $|\vartheta| > 1$.*

Proof. We go by induction on n. If $n = 1$ then $m = k$ and $g_1(X_k) \subseteq X_{3-k}$ by the properties of inductive pairs. Now assume $n \geq 1$ and that $\Phi(g_n \cdots g_1)(X_k) = g_n \cdots g_1(X_k) \subseteq X_{3-m}$, where $g_n \cdots g_1$ is a (m,k)-form; if $g_{n+1} \in G_{3-m} - J$ then $g_{n+1}(X_{3-m}) \subseteq X_m$ and thus $g_{n+1} \cdots g_1(X_k) \subseteq g_{n-1}(X_{3-m}) \subseteq X_m$. This proves the first statement.

We now prove the properness statement; suppose (X_1, X_2) is proper and $|\vartheta| > 1/$ Without loss of generality, assume the G_1-translates of X_1 do not cover X_2. We have two cases:

(a) If $g_1 \in G_1 \setminus J$, then $g_1(X_1) \subset X_2$ and so $g_n \cdots g_1(X_1) \subset X_{3-m}$.

(b) If $g_1 \in G_2 \setminus J$, then (here is where we use $|\vartheta| > 1$) $g_2 \in G_1 \setminus J$, so $g_2 g_1(X_1) \subset X_2$ by (a).

9.1.12 Theorem (Ping-pong lemma). *Suppose $G = \langle G_1, G_2 \rangle$ acts on X and suppose $G_1 > J < G_2$ is a shared subgroup. If there exists a proper interactive pair (X_1, X_2) for this data, then $G \simeq G_1 *_J G_2$.*

Proof. No normal form of length 1 can be the identity. Assume $|\vartheta| = n > 1$ for $\vartheta = g_n \cdots g_1$; then for each $m \in \{1,2\}$, $\Phi(\vartheta)(X_m)$ is properly contained in either X_1 or X_2 by Lemma 9.1.11 and so $\Phi(\vartheta) \neq 1$ (so $\vartheta \neq 1$).

9.2 Applications of amalgamated products to group theory

See [5] for this and further material.

9.2.1 Theorem (Higman, 1951). *There exists a finitely presented group G which is isomorphic to one of its proper factor groups.*

9.2.2 Theorem (Higman, Neumann, and Neumann, 1949). *Every countable group can be embedded in a 2-generator group.*

For a proof of Theorem 9.2.2 see [5, p. 105].

Proof of Theorem 9.2.1. Let $A = \langle a, s \,:\, sas^{-1} = a^2 \rangle$, $B = \langle b, t \,:\, tbt^{-1} = b^2 \rangle$, $G = A *_J B$ where J is the subgroup of $A * B$ generated by ab^{-1}. Let $H = \langle a \rangle$, $K = \langle b \rangle$, and let $\varphi : H \to K$ be the natural isomorphism sending $a \mapsto b$.

Define $\alpha : A \to G$ and $\beta : B \to G$ by

$$\begin{array}{cc} \alpha(a) = a^2 & \beta(b) = b^2 \\ \alpha(s) = s & \beta(t) = t; \end{array}$$

clearly α and β agree on H (in the sense that $\alpha(h) = \beta(\varphi(h))$ for all $h \in H$) and so there is a homomorphism $\mu : G \to G$ which restricts to $\alpha = \beta$ on the subgroup J (Lemma 9.1.6).

Note that μG includes $\{a^2, s, t\}$ and so μG contains $a = s^{-1}a^2 s$; hence μG includes $\{a, s, t\}$ which is a generating set for G, i.e. $\mu G = G$; hence $G \simeq G/\ker\mu$.

It remains to show that $\ker\mu \neq 1$: one can check easily that $sas^{-1}tb^{-1}t^{-1} \in \ker\mu$, and this is an alternating product of elements of A and B so is a normal form, hence not the identity.

We conclude by noting that the usual definition of the amalgamated product is as follows: if $A \leq G$ and $B \leq H$ are groups, and $\varphi : A \to B$ is an isomorphism, then we define

$$A *_\varphi B := \frac{G * H}{\overline{\{\varphi(a)a^{-1} \,:\, a \in A\}}}$$

(where $\overline{K}$ denotes the normal closure in Γ of some $K \leq \Gamma$).

This construction (and the construction following of *HNN-extensions*) is a special case of the principle of Bass-Serre theory in which various generalisations of free products are realised as fundamental groups of a 'graph of groups', see [41, section I.5].

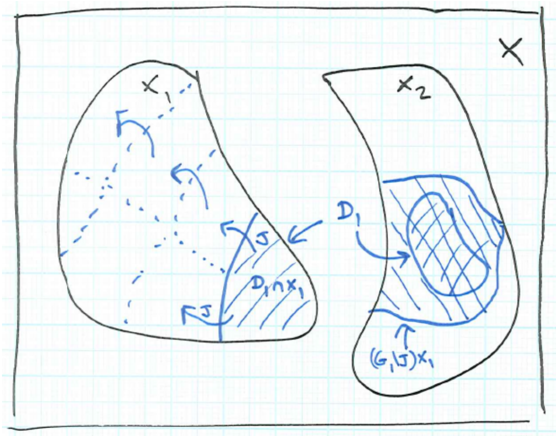


Figure 9.3: The interactive pair and one half of the packing D for Theorem 9.3.1.

9.3 The Klein combination theorem

Let G be a group acting on a space X, and let $Y \subseteq X$ be precisely invariant under $H \leq G$. We say that a fundamental set D for G is **maximal (with respect to** Y**)** if $D \cap Y$ is a fundamental set for the action of J on Y.

See Fig. 9.3 for the setup of the following.

9.3.1 Theorem. *Let $G_1 \geq J \leq G_2$ act discontinuously on X. Assume that there is an interactive pair of sets (X_1, X_2) for this system, and assume that for each $m \in \{1,2\}$ there exists a minimal fundamental set D_m for G_m such that for all $g \in G_m$,*

$$g(D_m \cap X_{3-m}) \subseteq X_{3-m}$$

Let $D = (D_1 \cap X_2) \cup (D_2 \cap X_1)$. Then D is a G-packing.

Proof. Without loss of generality, assume $x \in D_1 \cap X_2$ and that x is nontrivial.

Case that $g \in J$**.** If $g \in J$, then $g(x) \in X_2$ since X_2 is J-invariant; further, $g(x) \notin D_1$, since otherwise $x, g(x) \in D_1$ contradicting uniqueness of orbit representatives. Thus $g(x) \in X_2 \setminus D_1$ and thus $g(x) \notin D$.

Case that $g \notin J$**.** We proceed by induction on the length of g as a normal form. If $|g| = 1$, then we have two cases: either $g \in G_1 \setminus J$, or $g \in G_2 \setminus J$. In either case using the hypothesised properties of the X_m and D_m we have $g(x) \in X_1 \setminus D_1$.

Now suppose $g = g_n \cdots g_1$ is in normal form with $n > 1$; the inductive hypthesis takes the form 'if $h \in G$ is an m-form with $|h| < n$ then $\Phi(h)(x) \in Y_{3-m} \setminus D_m$'. Take $h = g_{n-1} \cdots g_1$. Case I: h is a 1-form. Then $g_{n-1} \in G_1 \setminus J$ so $h(x) \in X_2 \setminus D_1$, and $g_n \in G_2 \setminus K$ so $g_n(X_2) \setminus X_1$ i.e. $g(x) \in X_1$ and $g(x) \notin D_2$. Case II: h is a 2-form. Similar proof. ⛵

9.3.2 Theorem (Klein combination theorem)**.** *Let G_1, G_2 act freely and discontinuously on some open subset of a topological space X, and let $G = \langle G_1, G_2 \rangle$. Suppose for each m there exists a fundamental set D_m for G_m with $D_1 \cup D_2 = X$ and $D_1 \cap D_2 \neq \emptyset$. Then $G = G_1 * G_2$ and $D := D_1 \cap D_2$ is a G-packing.*

Proof. Assume G_1, G_2 nontrivial. Then both $D_1 \setminus D_2$ and $D_2 \setminus D_1$ are both nontrivial. Let $X_1 := D_1 \setminus D_2$ and $X_2 := D_2$. We show that (X_1, X_2) is a proper interactive pair for the system $G_1 \geq 1 \leq G_2$. If $g \in G_1 \setminus \{1\}$ then $gD_1 \cap D_1 \neq \emptyset$ since D_1 is fundamental for G_1; hence $gX_1 \subseteq gD_1 \subseteq D_2 = X_2$. If $g \in G_2 \setminus \{1\}$ then $gX_2 \subseteq X_1$ since $gX_2 = gD_2 \cap D_2 = \emptyset$. Further the pair is proper since if $x \in D = D_1 \cap D_2$ then $x \in X_2$ and $gx \notin X_1$ for any $g \in G_1$. We now use Theorem 9.1.12 to conclude that $G = G_1 * G_2$.

Now we show that D is a G-packing. Note that $D_m \cap X_m = X_m$ for each m, hence D_m is maximal; then apply Theorem 9.3.1. ☐

9.4 HNN-extensions

Motivation:

1. Given two isomorphic subgroups of a given group G, are they always conjugate in G? Answer: no, for instance if G is abelian then all conjugacy classes are trivial. However, there is some overgroup of G in which they are conjugate.

2. Algebraic topology: see e.g. [22, section 13.3] or [26, p. 93].

Here is the 'correct' definition:

9.4.1 Definition. Let G be a group, $A, B \leq G$, and $\varphi : A \to B$ an isomorphism. Suppose $\langle t \rangle$ is the free group on some symbol t unrelated to G. Then the **HNN-extension** of G relative to A and B with **stable letter** t is the group

$$G *_t := \langle G, t : \forall_{a \in A} t^{-1}at = \varphi(a) \rangle = \frac{G * \langle t \rangle}{\{a \in A : t^{-1}at(\varphi(a))^{-1}\}}.$$

Remark. It follows, e.g. from the lemma of Britton [41, I.5.2 theorem 11], that the quotient in the definition above does not collapse.

A more concrete definition is now given. Let G_0, G_1 and $J_1, J_2 \leq G_0$ be groups such that

1. $G_0 \cap G_1 = \emptyset$;

2. G_1 is the cyclic group on the symbol f;

3. the map $f_* : J_1 \to J_2$ defined by $f_* j = fjf^{-1}$ (the product taken in $\langle G_0, G_1 \rangle$) is an isomorphism.

9.4.2 Definition. A **normal form** is a word in $G_0 \cup G_1$ of the form $f^{\alpha_n} g_n \cdots f^{\alpha_1} g_1$, where

1. each $g_k \in G_0$;

2. for $k > 1$, $g_k \neq 1$;

3. all $\alpha_k \in \mathbb{Z}$, with only α_0 allowed to be 0;

4. if $\alpha_k < 0$ and $g_{k+1} \in J_1 \setminus 1$, then $\alpha_{k+1} < 0$;

5. if $\alpha_k > 0$ and $g_{k+1} \in J_2 \setminus 1$, then $\alpha_{k+1} > 0$.

Two such forms are **equivalent** if we may obtain one from the other by a finite sequence of insertions and deletions of words of the form $fjf^{-1}(f_*j)^{-1}$ or conjugates or inverses of these. Every word $f^{\alpha_n}g_n \cdots f^{\alpha_1}g_1$ is equivalent either to a normal form or to 1. The **HNN-extension** of G_0 by f, denoted $G_0 *_f$, is the set of normal forms (together with 1) modulo equivalence, with juxtaposition as the operation.

There exists a natural homomorphism $\Phi : G_0 *_f \to \langle G_0, G_1 \rangle$ defined by $\Phi(f^{\alpha_n}g_n \cdots f^{\alpha_1}g_1) = f^{\alpha_n}g_n \cdots f^{\alpha_1}g_1$.

Normal forms have unique representative in $G_0 *_f$ iff $J_1 = J_2 = 1$ (and in this case $G = G_0 * G_1$).

The **length** of a normal form is $|f^{\alpha_n}g_n \cdots f^{\alpha_1}g_1| := \sum_{m=1}^{n} |\alpha_m|$. This is well-defined by properties 4 and 5 of the definition above of normal forms. In addition, we say that $g = f^{\alpha_n}g_n \cdots f^{\alpha_1}g_1$ is **positive** ($g > 0$), **null** ($g \sim 0$), or **negative** ($g < 0$) if α_n is positive, 0, or negative respectively. Again by properties 4 and 5 of the definition, sign is well-defined.

9.4.3 Lemma. *Isomorphisms on G_0 and G_1 can be extended. More precisely, suppose $G = G_0 *_f$, $\tilde{G} = \tilde{G}_0 *_{\tilde{f}}$, and suppose f and $\tilde{f}$ respectively conjugate $J_1 \to J_2$ and $\tilde{J_1} \to \tilde{J_2}$. Given $\varphi_0 : G_0 \to \tilde{G}_0$ and $\varphi_1 : G_1 \to \tilde{G_1}$ both isomorphisms with $\varphi_1(f) = \tilde{f}$ and with $\varphi_0(J_m) = \tilde{J_m}$ ($m \in \{1,2\}$), if $\varphi_0 f_* \restriction_{J_1} - \tilde{f}_* \varphi_0 \restriction_{J_1}$, then we obtain an isomorphism $\varphi : G \to \tilde{G}$ which agrees with φ_0 and φ_1 on their respective domains.*

Chapter 10

Combinatorial group theory and algebraic geometry

In this chapter, we study combinatorial group theory from the point of view of Bass-Serre theory and algebraic geometry. Our primary source is [15]. We shall often reference [40] for representation theory, and [42] and [25] for algebraic geometry.

10.1 Representation theory

Let Π be a finitely generated group. Eventually, Π will be $\pi_1(M)$ for some 3-manifold M.

Recall that a **representation** of Π is a homomorphism $\rho : \Pi \to \mathrm{GL}(2, k)$, where k is a field. In this chapter, we will usually be interested in representations with image in $\mathrm{SL}(2, \mathbb{C})$; the set of all such representations will be denoted by $\mathrm{Rep}(\Pi)$. We say that two representations ρ_1 and ρ_2 are **equivalent** if there exists some $\alpha \in \mathrm{GL}(2, k)$ such that $\rho_1 = \alpha\rho_2\alpha^{-1}$.

Given a representation ρ, the **character** of ρ is the function $\chi_\rho : \Pi \to k$ defined by $\chi_\rho(g) = \mathrm{tr}\,\rho(g)$. Since tr is conjugacy invariant, $\rho_1 \sim \rho_2$ implies $\chi_{\rho_1} = \chi_{\rho_2}$. The converse of this statement is a standard fact in the finite group case (see for instance [40, corollary 2 to theorem 4 of chapter 2]) but is less standard in the finite generated group case, and requires additional qualification. Recall that a representation $\rho : \Pi \to \mathrm{GL}(n, k)$ is **irreducible** if the only subspaces of $\mathbb{R}^n$ irreducible under $\rho(\Pi)$ are 0 and k^n.

10.1.1 Theorem. *Let ρ_1 and ρ_2 be representations of Π into $\mathrm{GL}(n, \mathbb{C})$. If ρ_1 is irreducible, then $\rho_1 \sim \rho_2$ if and only if $\chi_{\rho_1} = \chi_{\rho_2}$.*

Let $\Pi = \langle g_1, \ldots, g_n \rangle$. Define

$$R(\Pi) := \{(\rho g_1, \ldots, \rho g_n) : \rho \in \mathrm{Rep}(\Pi)\}.$$

10.1.2 Lemma. *The set $R(\Pi)$ is an affine subvariety of $\mathbb{C}^{4n}$.*

For each choice of generators of Π, there is a natural bijective correspondence between the set $R(\Pi)$ defined with respect to that set of generators, and the set $\mathrm{Rep}(\Pi)$. Further, if $\{g_1, \ldots, g_n\}$ and $\{h_1, \ldots, h_m\}$ are generating sets for Π then the natural bijection

$$\{(\rho g_1, \ldots, \rho g_n) : \rho \in \mathrm{Rep}(\Pi)\} \to \{(\rho h_1, \ldots, \rho h_m) : \rho \in \mathrm{Rep}(\Pi)\}$$
$$(\rho g_1, \ldots, \rho g_n) \mapsto (\rho h_1, \ldots, \rho h_m)$$

is an isomorphism of varieties. Thus $R(\Pi)$, as an abstract variety, is well-defined independent of any choice of generating set.

Proof. Suppose that $\Pi = \langle g_1, ..., g_n \rangle$ and that W is the set of words in the generators which are trivial in Π. Suppose that $z \in R(\Pi)$ with respect to these generators; then $z = (\rho g_1, ..., \rho g_n)$ for some $\rho : \Pi \to \mathrm{SL}(2, \mathbb{C})$. Since ρ is a homomorphism, if $w \in W$ is of the form $w = g_{i_1}^{n_1} \cdots g_{i_r}^{n_r}$ then

$$(*) \qquad I = \rho(w) = (\rho g_{i_1})^{n_1} \cdots (\rho g_{i_r})^{n_r}.$$

In particular, we obtain for each $w \in W$ a set of four polynomial equations in the coefficients of elements of $\mathrm{SL}(2, \mathbb{C})^n$, one for each component of the matrix equation (*). Let S be the set of all of the equations in obtained in this way (so S contains four equations for each relation in G). If $z \in \mathrm{SL}(2, \mathbb{C})^n$ satisfies all of these relations, then we may define a representation $\rho \in \mathrm{Rep}(\Pi)$ by setting ρg_i to be the matrix corresponding to the ith matrix component of z. We have shown therefore that $R(\Pi)$ is the affine subvariety cut out of $\mathrm{SL}(2, \mathbb{C})^n$ by the polynomials of S.

To show that the bijection in the statement is an isomorphism of varieties, it suffices to note that the equations for the g_i in terms of the h_j are polynomial in the entries of the matrices in the image of the representation. ∎

Due to this lemma, it is a valid abuse of notation to identify points in $R(\Pi)$ with representations of Π.

10.1.3 Proposition. *Let $V \subseteq R(\Pi)$ be an irreducible component and let $\rho \in V$. If $\sigma \in R(\Pi)$ is equivalent to ρ in $\mathrm{SL}(2, \mathbb{C})$, then $\sigma \in V$.*

Proof. Consider the morphism of varieties

$$\begin{aligned} f : V \times \mathrm{SL}(2, \mathbb{C}) &\to R(\Pi) \\ (x, \alpha) &\mapsto (\alpha x_1 \alpha^{-1}, ..., \alpha x_n \alpha^{-1}); \end{aligned}$$

since $\mathrm{SL}(2, \mathbb{C})$ is irreducible, $V \times \mathrm{SL}(2, \mathbb{C})$ is irreducible and since f is a morphism the image $f(V \times \mathrm{SL}(2, \mathbb{C}))$ is an irreducible subvariety of $R(\Pi)$. Hence there exists a component V' of $R(\Pi)$ such that $f(V \times \mathrm{SL}(2, \mathbb{C})) \subseteq V'$. On the other hand, $V = f(V \times \{1\})$ and so $V \subseteq V'$. It follows that $V = V'$ since both V and V' are irreducible components of $R(\Pi)$.

Suppose now that σ is equivalent to ρ; then there exists some $\alpha \in \mathrm{SL}(2, \mathbb{C})$ such that $\sigma = \alpha \rho \alpha^{-1}$. Thus for each i, $\sigma g_i = \alpha x_i \alpha^{-1}$ where $x = (x_1, ..., x_n) \in V$ is the point corresponding to ρ. In particular, $(\sigma g_1, ..., \sigma g_n) \in V$ by the previous paragraph, as desired. ∎

If k is a field, we say that a representation $\rho : \Pi \to \mathrm{GL}(n, k)$ is **absolutely irreducible** if it is irreducible when considered as a representation into $\mathrm{GL}(n, \bar{k})$.

10.1.4 Proposition. *Let k be a field of characteristic 0, and let $\rho : \Pi \to \mathrm{SL}(2, k)$ be a representation with non-abelian image. Then the following are equivalent:*

1. *ρ is reducible;*
2. *ρ is reducible over $\bar{k}$;*
3. *for all $c \in [\Pi, \Pi]$, $\chi_\rho(c) = 2$;*
4. *the group $\rho([\Pi, \Pi])$ has a unique invariant subspace $L \leq k^2$ of dimension 1.*

Further, the implication (1) $\implies$ (3) *does not require $\rho(\Pi)$ to be non-abelian.*

Proof. The implication (1) $\implies$ (2) is trivial.

To show (2) $\implies$ (3), suppose that ρ is reducible over $\bar{k}$. Then there exists a representation σ of Π with image consisting of upper triangular matrices such that $\sigma \sim \rho$. Let $A, B \in \sigma(\Pi)$. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $B = \begin{bmatrix} x & y \\ 0 & w \end{bmatrix}$, then

$$[A, B] = ABA^{-1}B^{-1} = \begin{bmatrix} 1 & (dw)^{-1}(bw - bx + ay - dy) \\ 0 & 1 \end{bmatrix}$$

so $\operatorname{tr}[A, B] = 2$ as desired.

We next show that (3) $\implies$ (4). Since $\rho(\Pi)$ is non-abelian, there exists some $c \in [\Pi, \Pi]$ such that $\rho(c) \neq 1$. Since $\chi_\rho(c) = 2$ we have $\operatorname{tr} \rho(c) = 2$ and $\det \rho(c) = 1$ and so the eigenvalues of $\rho(c)$ are the solutions in λ of the characteristic equation

$$0 = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2;$$

i.e. $\rho(c)$ has a single eigenvalue, $\lambda = 1$. If the eigenspace of this eigenvalue had dimension 2, then $\rho(c)$ would fix the entirety of k^2 and so $\rho(c) = 1$, which is impossible by assumption; thus the eigenspace of λ has dimension 1 and $\rho(c)$ has a unique invariant subspace L of dimension 1.

It remains to show that this subspace L is invariant under every element of $[\Pi, \Pi]$. Suppose that there exists $c' \in [\Pi, \Pi]$ such that $\rho(c')L \not\subseteq L$. Then $\rho(c') \neq 1$ and by the above argument there exists a unique subspace L' of dimension 1 left invariant by $\rho(c')$ distinct from L. Pick $l \in L$ and $l' \in L'$; then $\{l, l'\}$ is a basis for k^2 and with respect to this basis we have

$$\rho(c) = \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix}, \qquad \rho(c') = \begin{bmatrix} 1 & 0 \\ \beta & 1 \end{bmatrix}$$

where $\alpha, \beta \neq 0$; in particular, $\operatorname{tr} \rho(cc') = 2 + \alpha\beta \neq 2$ which provides a contradiction and establishes that (3) $\implies$ (4).

Finally we show that (4) $\implies$ (1). Suppose L is the unique dimension 1 subspace left invariant by $\rho([\Pi, \Pi])$. Since $\rho([\Pi, \Pi])$ is normal in $\rho(\Pi)$, L is invariant under $\rho(\Pi)$: indeed, if $L = kl$ then pick a nontrivial $k \in \rho(\Pi')$; for all $g \in \rho(\Pi)$, we may compute $gl = g(g^{-1}kg)l = k(gl)$, so gl is fixed by k and hence by uniqueness of the fixed line we have $gl \in L$, so $gL \subseteq L$. In particular, $\rho(\Pi)$ fixes a nontrivial subspace and so ρ is reducible. ∎

10.1.5 Corollary. *If K is an algebraically closed field of characteristic 0, then $\rho : \Pi \to \mathrm{SL}(2, \mathbb{C})$ is reducible iff $\chi_\rho(c) = 2$ for all $c \in [\Pi, \Pi]$.*

Proof. Observe that the implication (1) $\implies$ (3) of Proposition 10.1.4 did not use the assumption that $\rho(\Pi)$ was non-abelian. This supplies one direction of the corollary. For the converse, if $\rho(\Pi)$ is non-abelian then the result follows directly from the theorem.

Otherwise, suppose that $\rho(\Pi)$ is abelian. Recall that every irreducible representation of an abelian group over such a field K is of degree 1 (this follows from Schur's lemma, which in the version stated in [40] does not require $|\Pi| < \infty$: if $g \in \Pi$ then $v \mapsto (\rho g)v$ is Π-linear by commutativity and so by Schur's lemma the map is given by a multiple of the identity matrix, in particular ρg is diagonal for all $g \in \Pi$ and so leaves Ke_i invariant for each standard basis vector e_i of K^2). In particular, $\rho(\Pi)$, being a degree 1 representation in a dimension 2 vector space, is reducible. ∎

Remark. Compare Proposition 10.1.4 and Corollary 10.1.5 with the results of Section 2.4 above.

10.2 Curves of representations and character varieties

We use some standard facts about projective completions and resolution of singularities of algebraic curves; for instance, see [25, corollary I.6.11] or [42, corollary to theorem 2.23].

10.2.1 Theorem. *Let C be an affine algebraic curve over a field k. Then there exists a projective curve $\bar{C}$ and a nonsingular projective curve $\tilde{C}$ such that $\mathbf{K}(\tilde{C}) \simeq \mathbf{K}(\bar{C}) \simeq \mathbf{K}(C)$.*

If C and D are affine curves and if $f : C \to D$ is a rational map, then there exists a unique regular map $\tilde{f} : \tilde{C} \to \tilde{D}$.

Suppose that C is an affine curve in $R(\Pi)$. We define a representation $P : \Pi \to \mathrm{SL}(2, k)$ where $k = \mathbf{K}(\tilde{C})$, called the **canonical representation** with respect to C. If $g \in \Pi$ is fixed, define $a, b, c, d : C \to \mathbb{C}$ (all depending on g) by the system of equations

$$\begin{bmatrix} a(\rho) & b(\rho) \\ c(\rho) & d(\rho) \end{bmatrix} := \rho(g);$$

we then define $P(g)$ to be the matrix

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

10.2.2 Lemma. *If C contains an irreducible representation, then P is absolutely irreducible.*

Proof. If P is reducible over $\bar{k}$, then by (2) $\implies$ (3) of Proposition 10.1.4, $\operatorname{tr} P(c) = 2$ for all $c \in [\Pi, \Pi]$. By definition of P, $\operatorname{tr} P(c) = \tilde{f}_c$ where $f_c(\rho) = a(\rho) + d(\rho) = \operatorname{tr} \rho(c)$ for all $\rho \in C$. Thus $\operatorname{tr} \rho(c) = 2$ for all $c \in [\Pi, \Pi]$ and all $\rho \in C$. Now by Corollary 10.1.5 (noting that such ρ are representations over $\mathbb{C}$ not k) we conclude that each $\rho \in C$ is reducible.

For each $g \in \Pi$, define a regular map $\tau_g : R(\Pi) \to \mathbb{C}$ by $\tau_g(\rho) := \chi_\rho(g)$. Let T be the subring of $\mathbf{K}(R(\Pi))$ generated by the set $\{\tau_g : g \in \Pi\}$.

10.2.3 Lemma. *For all $g, h \in \Pi$, $\tau_g \tau_h = \tau_{gh} + \tau_{gh^{-1}}$.*

10.2.4 Proposition. *The ring T is finitely generated.*

Proof. Let $T_0 \subseteq T$ be the subring generated by all the maps $\tau_{g_{i_1} \cdots g_{i_r}}$, where $i_1, \dots, i_r$ are r distinct elements of $\{1, \dots, n\}$. We will show that $T_0 = T$.

Suppose first that $g = g_{i_1}^{m_1} \cdots g_{i_r}^{m_r}$ where $i_1, \dots, i_r \in \mathbb{N}$ are distinct and $m_1, \dots, m_r \in \mathbb{Z}$. We show that $\tau_g \in T_0$ by induction on the quantity

$$\nu = \sum_{i=1}^{r} K_i$$

where

$$K_i = \begin{cases} -m_i & m_i \leq 0, \\ m_i - 1 & m_i > 0. \end{cases}$$

If $\nu = 0$, then all of the m_i are 0 or 1 and so g is a generator of T_0. If $\nu > 0$, then we may assume that $m_r \neq 0$ (otherwise $g_{i_r}^{m_r} = 1$ and so we can simply cease to write it). If $m_r = 1$, then consider

$$g_{i_r} g g_{i_r}^{-1} = g_{i_r} g_{i_1}^{m_1} \cdots g_{i_{r-1}}^{m_{r-1}}$$

which has the same value of ν as g; repeating the same process and using the fact that not all of the m_i are 0 or 1, we may assume (after repeated conjugations, which preserve τ.) that $m_r \neq 1$ too.

If $m_r < 0$, then by Lemma 10.2.3 we have

$$\tau_{gg_{i_r}} \tau_{g_{i_r}^{-1}} = \tau_g + \tau_{gg_{i_r}^2};$$

by induction, $\tau_{gg_{i_r}}$ and $\tau_{gg_{i_r}^2}$ lie in T_0; and by definition, $\tau_{g_{i_r}^{-1}} \in T_0$; so $\tau_g \in T_0$. If $m_r > 1$ then a similar argument shows that $\tau_g \in T_0$.

Now suppose $g \in \Pi$ is arbitrary; write $g = g_{i_1}^{m_1} \cdots g_{i_r}^{m_r}$ and induct on r. We may assume that the i_k are not all distinct; by a similar conjugation argument as above, we may assume that there is some $s < r$ such that $i_s = i_r$. Now set $V = g_{i_1}^{m_1} \cdots g_{i_s}^{m_s}$ and $W = g_{i_{s+1}}^{m_{s+1}} \cdots g_{i_r}^{m_r}$. By Lemma 10.2.3, $\tau_g = \tau_{VW} = \tau_V \tau_W - \tau_V \tau_{W^{-1}}$; but all of τ_V, τ_W, and $\tau_V \tau_{W^{-1}}$ lie in T_0 by induction; and hence $\tau_g \in T_0$.

Since T is finitely generated, there exist $\gamma_1, ..., \gamma_m \in \Pi$ such that $T = \langle \tau_{\gamma_i} : 1 \leq i \leq m \rangle$. Define a map $t : R(\Pi) \to \mathbb{C}^m$ by

$$t(\rho) := (\tau_{\gamma_1}(\rho), ..., \tau_{\gamma_m}(\rho)).$$

As in the case of $R(\Pi)$, there is a natural correspondence between the points of $X(\Pi) := t(R(\Pi))$ and the characters of representations of Π in $\mathrm{SL}(2, \mathbb{C})$.

Our immediate goal is now the proof that $X(\Pi)$ is a closed subvariety of $\mathbb{C}^m$.

10.2.5 Lemma. *The set of reducible representations* $\Pi \to \mathrm{SL}(2, \mathbb{C})$ *is of the form* $t^{-1}(V)$ *for some closed subvariety* $V \subseteq \mathbb{C}^m$.

Proof. By Corollary 10.1.5, a point $\rho \in R(\Pi)$ is reducible iff $\tau_c(\rho) = 2$ for all $c \in [\Pi, \Pi]$. Since T is generated by the τ_{γ_i}, the function $\tau_c(\cdot)$ is of the form $f(t(\cdot))$ for some f a polynomial function with integral coefficients. This exhibits the set of reducible representations as the inverse image of $\mathbf{Z}(f - 2)$ via t.

10.2.6 Lemma. *Let A be a principal ideal domain, let $F = \mathrm{Frac}\, A$, and let $P : \Pi \to \mathrm{GL}(n, F)$ be an absolutely irreducible representation for some $n > 0$. If $\chi_P(\Pi) \subseteq A$ then P is equivalent to a representation with image in* $\mathrm{GL}(n, A)$.

Proof.

Appendix A

Algebraic topology

We review some basic concepts from algebraic topology, such as are contained in many excellent books like [33, 8, 7].

A.1 Homotopy

Let X be a topological space, and fix some $x_0 \in X$. A continuous map $f : [0,1] \to X$ is a **loop based at** x_0 if $f(0) = x_0 = f(1)$.

Let f and g be loops based at x_0; we define their **concatenation** $f * g : [0,1] \to X$ to be the map

$$t \mapsto \begin{cases} f(2t) & t \leq 1/2 \\ g(2t-1) & t \geq 1/2. \end{cases}$$

In addition, we say that f and g are **homotopic** and write $f \simeq g$ if there is a continuous map $F : [0,1] \times [0,1] \to X$ (the **homotopy**) such that

- $F(x,0) = f(x)$ and $F(x,1) = g(x)$ for all $x \in [0,1]$;
- $F(x_0, t) = x_0$ for all $t \in [0,1]$.

More generally, a **homotopy** is a continuous map $W \times [0,1] \to X$ for W an arbitrary topological space.

The following elementary fact is easy to prove (take the concatenation of the two homotopy maps in the obvious way):

A.1.1 Lemma. *If $f_1 \simeq f_2$ and $g_1 \simeq g_2$, then $(f_1 * g_1) \simeq (f_2 * g_2)$.*

This shows, in particular, that concatenation is well-defined on the equivalence classes of loops based at x_0. It is easy to see that:

- if $f, g, h : [0,1] \to X$ are loops then $(f * g) * h \simeq f * (g * h)$;
- the constant map $\iota : [0,1] \ni t \mapsto x_0 \in X$ satisfies $\iota * f \simeq f \simeq f * \iota$ for all loops f; and
- for a loop f we may define $\hat{f} : [0,1] \to X$ by $\hat{f}(t) := f(1-t)$ for all $t \in [0,1]$, and that $f * \hat{f} = \iota = \hat{f} * f$.

Thus the set of all equivalence classes under homotopy of loops based at x_0 forms a group under the operation $*$; we denote this group, called the **first homotopy group** or the **fundamental group**, by the symbol $\pi_1(X, x_0)$. Observe that, *a priori*, the group depends on x_0. In fact, if X is path-connected then it is independent of the basepoint up to isomorphism (but *not* up to natural isomorphism).

A.1.2 Lemma. *If X is path-connected and $x_0, x_1 \in X$ then $\pi_1(X, x_0) \simeq \pi_1(X, x_1)$.*

To prove the lemma, it suffices to observe that if $\alpha : [0,1] \to X$ is a path such that $\alpha(0) = x_0$ and $\alpha(1) = x_1$ then the map $\pi_1(X, x_0) \to \pi_1(X, x_1)$ defined on equivalence class representatives by $f \mapsto \alpha f \alpha^{-1}$ is well-defined up to homotopy and is an isomorphism of groups.

We give some standard examples whose proofs may be found in any algebraic topology textbook such as those listed above.

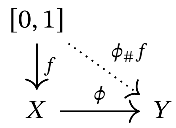
A.1.3 Example.

1. $\pi_1(S^1) \simeq \mathbb{Z}$
2. $\pi_1(S^n) \simeq \pi_1(\mathbb{R}^2) \simeq 1$ for $n \geq 2$
3. $\pi_1(T_{1,0}) \simeq \mathbb{Z} \oplus \mathbb{Z}$ ($T_{1,0} \simeq S^1 \times S^1$ is the torus with a single 'hole')
4. $\pi_1(T_{1,1}) \simeq \mathbb{Z} * \mathbb{Z}$ ($T_{1,1}$ is the torus with a single 'hole' and a single puncture)
5. $\pi_1(T_{1,r}) \simeq *^{r+1}\, \mathbb{Z}$ ($T_{1,r}$ is the torus with a single 'hole' and r punctures)

Finally, observe that if X and Y are topological spaces, and if $\phi : X \to Y$ is a continuous map, then for each choice of x_0 we have a natural map

$$\phi_\# : \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$$

via the diagram



(that is, $\phi_\# f := \phi \circ f$ for all $f \in \pi_1(X, x_0)$). It is easy to check that the maps sending $X \mapsto \pi_1(X, x_0)$ and $\phi \mapsto \phi_\#$ define a functor (i.e. $\phi_\# \circ \psi_\# = (\phi\psi)_\#$).

A.2 Covering spaces and deck transformations

The definition of a covering space is motivated by the generalisation of various classical examples:

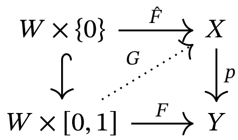
A.2.1 Example.

1. The 'spiral covering' of S^1 by $\mathbb{R}$ (namely, $f : \mathbb{R} \to S^1 \subseteq \mathbb{C}$ defined by $f(t) = \exp(it)$)
2. The covering of the torus $T_{1,0}$ by $\mathbb{R}^2$ (namely, the projection map $\mathbb{R}^2 \to \mathbb{R}^2/\mathbb{Z}^2$).

The formal definition follows.

A.2.2 Definition. Let X and Y be topological spaces which are Hausdorff, path-connected, and locally path-connected. A continuous map $p : X \to Y$ is a **covering map** if each point $y \in Y$ has a path-connected neighbourhood U such that $p^{-1}(U)$ is a nonempty disjoint union $\bigcup_{\alpha \in A} U_\alpha$ with the property that $p\restriction_{U_\alpha} : U_\alpha \to U$ is a homeomorphism for each $\alpha \in A$.

A.2.3 Theorem (Covering Homotopy Theorem)**.** *Let W be a locally connected topological space and let $p : X \to Y$ be a covering map. Let $F : W \times [0,1] \to Y$ be a homotopy, and let $\hat{F} : W \times \{0\} \to X$ be any map such that $(p \circ \hat{F})(w, 0) = F(w, 0)$ (such a map exists since p is surjective). Then there is a unique homotopy $G : W \times [0,1] \to X$ making the following diagram commute:*



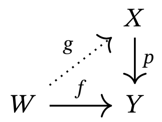
Moreover, if $A \subseteq W$ is such that $F(a,t)$ is independent of t for all $a \in A$, then $G(a,t)$ is also independent of t for all $a \in A$.

A.2.4 Corollary. *Let $p : X \to Y$ be a covering map and let $x_0 \in X$ be fixed. Then the map $p_\# : \pi_1(X, x_0) \to \pi_1(Y, p(x_0))$ is a monomorphism whose image consists of the classes of loops at $p(x_0)$ in Y which lift to loops at x_0 in X.*

Proof. Suppose $f \in \ker \pi_1(X, x_0)$; then there exists a homotopy $F : [0,1] \times [0,1] \to Y$ with $F(x,0) = (p_\# f)(x) = pf(x)$ and $F(x,1) = p(x_0)$ for all $x \in [0,1]$. Let $\hat{F}$ and G be the lifts of F as in the theorem ($\hat{F}$ chosen arbitrarily); it will suffice to check that G is a homotopy such that $G(x,0) = f(x)$ and $G(x,1) = x_0$ for all $x \in [0,1]$. But this follows immediately from the diagram.

If the image $p_\# \pi_1(X, x_0)$ is normal in $\pi_1(Y, p(x_0))$, then we say that the covering p is a **normal covering** or a **regular covering** (with respect to the basepoint x_0).

A.2.5 Theorem. *Let $p : X \to Y$ be a covering map, with $x_0 \in X$ distinguished. If W is a path-connected and locally path-connected topological space, $w_0 \in W$, and $f : W \to Y$ satisfies $f(w_0) = p(x_0)$, then there exists a map $g : W \to X$ such that the following diagram exists*



if and only if $f_\# \pi_1(W, w_0) \subseteq p_\# \pi_1(X, x_0)$ in $\pi_1(Y, p(x_0))$.

A.2.6 Proposition. *Let W be connected, $p : X \to Y$ a covering map, and $f : W \to Y$ continuous. Let $g_1, g_2 : W \to X$ be lifts of f. If $g_1(w) = g_2(w)$ for some $w \in W$, then $g_1 = g_2$.*

There is a bijective correspondence

(A.2.7) $$\{\text{right cosets of } p_\# \pi_1(X, x_0) \text{ in } \pi_1(Y, p(x_0))\} \leftrightarrow \text{the fibre } p^{-1}(p(x_0))$$

which comes from the right action of $J = \pi_1(Y, p(x_0))$ on $F = p^{-1}(p(x_0))$ defined for $\alpha \in \pi_1(Y, y_0)$ and $x \in F$ by

$$x \cdot \alpha :=:= g(1)$$

where g is defined as a lift of some representative of α in Y to X via Theorem A.2.5.

We now consider maps which move the covering space but not the base space; for instance, consider the map $t \mapsto t + 2\pi$ on $\mathbb{R}$, which behaves well with the covering $t \mapsto \exp(ti)$ of S^1 in Example A.2.1.

A.2.8 Definition. Let $p : X \to Y$ be a covering map. A homeomorphism $D : X \to X$ is a **deck transformation** or an **automorphism** of the covering if $p \circ D = p$.

The deck transformations under conjugation form a group, $\mathrm{Aut}(p)$. The deck transformations have a natural action on the fibres of p

A.2.9 Lemma. *The covering $p : X \to Y$ is normal with respect to the basepoint x_0 iff $\mathrm{Aut}(p)$ acts transitively on the fibre $p^{-1}p(x_0)$.*

A.2.10 Theorem. *Let $p : X \to Y$ be a covering map, and let $N(H)$ denote the normaliser of a subgroup $H \leq \pi_1(Y, p(x_0))$. There is an isomorphism*

$$\mathrm{Aut}(p) \simeq \frac{N(p_\# \pi_1(X, x_0))}{p_\#(\pi_1(X, x_0))}$$

which is induced by the epimorphism $\Theta : N(\mathrm{Stab}\, x_0) \to \mathrm{Aut}(p)$ defined by $\alpha \mapsto D_\alpha$, where $D_\alpha \in \mathrm{Aut}(p)$ is the unique deck transformation such that $D_\alpha(x_0) = x_0 \cdot \alpha$.

A.2.11 Corollary. *If $p : X \to Y$ is a normal covering map with respect to the basepoint x_0, then $\mathrm{Aut}(p) \simeq \pi_1(Y, p(x_0))/\pi_1(X, x_0)$.*

A.2.12 Corollary. *If $p : X \to Y$ is a covering map with X simply connected, then $\mathrm{Aut}(p) \simeq \pi_1(Y, p(x_0))$ $(x_0 \in X)$.*

A.3 Freely discontinuous actions

We repeat the opening definition of Section 3.1:

A.3.1 Definition. Let X be a topological space, and let G be a group with an action as a group of homeomorphisms on X. The action is said to be **freely discontinuous** (or **properly discontinuous**, or a **covering space action**) on X if, for every $x \in X$, there exists a neighbourhood $U \ni x$ (called a **nice neighbourhood**) such that $gU \cap U = \emptyset$ for all $g \in G$ nontrivial.

A.3.2 Proposition. *If G acts freely discontinuously on a path-connected and locally path-connected Hausdorff space X, then $p : X \to X/G$ is a regular covering map with $\mathrm{Aut}(p) = G$.*

Proof. Let $U \subseteq X$ be a nice neighbourhood of some $x \in X$; we may choose U to be path-connected. Set $U^* = p(U)$; this set is open since the projection is an open map. The connected components of U^* are the sets gU for $g \in G$. The maps $gU \to U^*$ are continuous, open, injective, and surjective, so are homeomorphisms; thus p is a covering map and clearly G acts as a subgroup of $\mathrm{Aut}(p)$ by left multiplication. We may apply Proposition A.2.6 to conclude that there are no others.

A.3.3 Corollary. *If X is simply connected and locally path-connected, and if G acts freely discontinuously on X, then $\pi_1(X/G) \simeq G$.*

Proof. Combine the above proposition with Corollary A.2.12.

Appendix B

Measures

We follow the terminology of [39].

A collection $\mathfrak{M}$ of subsets of a set X is a **σ-algebra** if it has the following properties:

1. $X \in \mathfrak{M}$;

2. If $A \in \mathfrak{M}$ then $X \setminus A \in \mathcal{M}$;

3. If $A_n \in \mathfrak{M}$ for $n \in \mathbb{N}$ and $A = \cup_{n=1}^{\infty} A_n$ then $A \in \mathfrak{M}$.

A set X equipped with a σ-algebra $\mathcal{M}$ is a **measurable space**. The members of $\mathcal{M}$ are the **measurable sets**. A map $f : X \to Y$, where X is a measurable space and Y is a topological space, is **measurable** if $f^{-1}(V)$ is a measurable set for each $V \subseteq Y$ open.

B.0.1 Theorem. *If $\mathcal{F}$ is any family of subsets of X, then there exists a smallest σ-algebra $\mathfrak{M}$ in X such that $\mathcal{F} \subseteq \mathfrak{M}$.*

B.0.2 Example. Let X be a topological space. By Theorem B.0.1, there exisrs a smallest σ-algebra $\mathfrak{B}$ in X such that every open set of X lies in $\mathfrak{B}$. The members of this algebra are the **Borel sets** on X.

If $\mathfrak{M}$ is a σ-algebra then a function $\mu : \mathfrak{M} \to [0, \infty]$ is a **measure** if it is countably additive: that is, if $\{A_n\}_{n\in\mathbb{N}}$ is a collection of disjoint members of $\mathfrak{M}$, then

$$\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n).$$

B.0.3 Example. Let $\mathfrak{M}$ be the collection of all subsets of some set X; then for $x \in X$ and $m \in [0, \infty]$, the function

$$\begin{aligned} \mathfrak{a} : \mathfrak{M} &\to [0, \infty] \\ E &\mapsto \begin{cases} m & x \in E \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

is a measure on $\mathfrak{M}$, the **atomic measure** at x with mass m.

A measure μ defined on the σ-algebra $\mathfrak{B}$ of Borel sets on a locally compact Hausdorff space X is called a **Borel measure** on X. We say that μ is **regular** if

- for every $E \in \mathfrak{B}$, we have

$$\mu(U) = \inf\{\mu(V) : V \text{ open and containing } E\};$$

- for every $E \in \mathfrak{B}$ either open in X or with finite measure, we have

$$\mu(U) = \sup\{\mu(K) : K \text{ compact and contained in } E\}.$$

Bibliography

[1] Andrew (https://math.stackexchange.com/users/30881/andrew). *Continuous Deformation Of Punctured Torus*. Mathematics Stack Exchange. Aug. 28, 2012. URL: `https://math.stackexchange.com/q/187617` (cit. on p. 74).

[2] Lars V. Ahlfors. *Complex analysis*. 3rd ed. McGraw Hill, 1979 (cit. on pp. 8, 53).

[3] Lars V. Ahlfors. "Correction to "Finitely generated Kleinian groups"". In: *American Journal of Mathematics* 87.3 (1965), p. 759 (cit. on p. 54).

[4] Lars V. Ahlfors. "Finitely generated Kleinian groups". In: *American Journal of Mathematics* 86.2 (1964), pp. 413–429 (cit. on p. 54).

[5] Gilbert Baumslag. *Topics in combinatorial group theory*. Lectures in Mathematics ETH Zürich. Birkhaüser, 1993 (cit. on p. 105).

[6] Alan F. Beardon. *The geometry of discrete groups*. Graduate Texts in Mathematics 91. Springer-Verlag, 1983 (cit. on pp. v, 2, 7, 8, 11, 13, 15, 24, 25, 31, 37, 93, 94).

[7] Raoul Bott and Loring W. Tu. *Differential forms in algebraic topology*. Graduate Texts in Mathematics 82. Springer, 1982 (cit. on p. 115).

[8] Glen E. Bredon. *Topology and geometry*. Graduate Texts in Mathematics 139. Springer, 1993 (cit. on p. 115).

[9] Gerhard Burde and Heiner Zieschang. *Knots*. 2nd ed. Studies in Mathematics 5. de Gruyer, 2003 (cit. on p. 74).

[10] Richard D. Canary, David Epstein, and Peter L. Green. "Fundamentals of hyperbolic geometry: Selected expositions". In: ed. by Richard D. Canary, David Epstein, and Albert Marden. LMS Lecture Note Series 328. Cambridge University Press, 2006 (cit. on p. 79).

[11] Richard D. Canary, David Epstein, and Peter L. Green. "Notes on notes of Thurston". In: *Analytical and geometric aspects of hyperbolic space*. Ed. by David Epstein. LMS Lecture Note Series 111. Cambridge University Press, 1987 (cit. on p. 70).

[12] Richard D. Canary, David Epstein, and Albert Marden, eds. *Fundamentals of hyperbolic geometry: Selected expositions*. LMS Lecture Note Series 328. Cambridge University Press, 2006 (cit. on p. 61).

[13] H. S. M. Coxeter. *Introduction to geometry*. 2nd ed. John Wiley & Sons, 1969 (cit. on p. 7).

[14] H. S. M. Coxeter and S. L. Greitzer. *Geometry revisited*. New Mathematical Library 19. MAA, 1967 (cit. on pp. 2, 5).

[15] Marc Culler and Peter B. Shalen. "Varieties of Group Representations and Splittings of 3-Manifolds". In: *Annals of Mathematics*. 2nd ser. 117.1 (1983), pp. 109–146 (cit. on p. 109).

[16] Simon Donoldson. *Riemann surfaces*. Oxford Graduate Texts in Mathematics 22. Oxford University Press, 2011 (cit. on p. 51).

[17] David Epstein, ed. *Analytical and geometric aspects of hyperbolic space*. LMS Lecture Note Series 111. Cambridge University Press, 1987 (cit. on p. 61).

[18] David Epstein and Albert Marden. "Convex hulls in hyperbolic space, and a theorem of Sullivan, and measured pleated surfaces". In: *Analytical and geometric aspects of hyperbolic Space*. Ed. by David Epstein. LMS Lecture Note Series 111. Cambridge University Press, 1987 (cit. on pp. 61, 70, 71).

[19] Günter Ewald. *Combinatorial convexity and algebraic geometry*. Graduate texts in mathematics 168. Springer, 1996 (cit. on p. 61).

[20] Otto Forster. *Lectures on Riemann surfaces*. Graduate Texts in Mathematics 81. Springer, 1981 (cit. on p. 51).

[21] George K. Francis. *A topological picturebook*. Springer, 1987 (cit. on pp. 66, 75).

[22] Stefan Friedl. *Algebraic Topology I - III*. Lecture notes. URL: `https://www.uni-regensburg.de/Fakultaeten/nat_Fak_I/friedl/papers/friedl-algebraic-topology.pdf` (cit. on p. 107).

[23] L. Greenberg. "On a theorem of Ahlfors and on conjugate subgroups of Kleinian groups". In: *American Journal of Mathematics* 89 (1967), pp. 56–68 (cit. on p. 54).

[24] Charlie Gunn, Delle Maxwell, and D. B. A Epstein. *Not Knot*. VHS tape. 1991. URL: `https://www.youtube.com/watch?v=4aN6vX7qXPQ` (cit. on p. 86).

[25] Robin Hartshorne. *Algebraic geometry*. Graduate texts in mathematics 52. Springer, 1977 (cit. on pp. 109, 112).

[26] Allen Hatcher. *Algebraic topology*. 2001. URL: `https://pi.math.cornell.edu/~hatcher/AT/AT.pdf` (visited on 2020) (cit. on p. 107).

[27] David C. Kay and Eugene W. Womble. "Axiomatic convexity theory and relationships between the Carathéodory, Helly, and Radon numbers". In: *Pacific Journal of Mathematics* 38.2 (1971), pp. 471–485 (cit. on p. 61).

[28] Linda Keen and Caroline Series. "Continuity of convex hull boundaries". In: *Pacific Journal of Mathematics* 168.1 (1995), pp. 183–206 (cit. on pp. 89, 91).

[29] Linda Keen and Caroline Series. "The Riley slice of Schottky space". In: *Proceedings of the London Mathematics Society* 3.1 (69 1994), pp. 72–90 (cit. on pp. 48, 57).

[30] S. L. Krushka, B. N. Apanasov, and N. A. Gusevski. *Kleinian groups and uniformization in examples and problems*. Ed. by Bernard Maskit. Trans. by H. H. McFaden. Translations of Mathematical Monographs 62. American Mathematical Society, 1986 (cit. on p. 54).

[31] Ravi S. Kulkarni and Peter B. Shalen. "On Ahlfors' finiteness theorem". In: *Advances in Mathematics* 76 (2 1989), pp. 155–169 (cit. on p. 54).

[32] John M. Lee. *Introduction to Riemannian manifolds*. 2nd ed. Graduate Texts in Mathematics 176. Springer, 2018 (cit. on p. 12).

[33] John M. Lee. *Introduction to topological manifolds*. 2nd ed. Graduate Texts in Mathematics 202. Springer, 2011 (cit. on pp. 30, 52, 73, 102, 115).

[34] Bernard Maskit. *Kleinian groups*. Grundlehren der mathematischen Wissenshaften 287. Springer-Verlag, 1987 (cit. on pp. v, 15, 24, 25, 31, 32, 48, 83).

[35] Katsuhiko Matsuzaki and Masahiko Taniguchi. *Hyperbolic manifolds and Kleinian groups*. Oxford University Press, 1998 (cit. on p. 87).

[36] Toshitsune Miyake. *Modular forms*. Monographs in Mathematics. Springer, 1989 (cit. on p. 32).

[37] R.L. Moore. “Concerning upper semi-continuous collections of continua”. In: *Transactions of the American Mathematical Society* 27 (1925), pp. 416–428 (cit. on p. 89).

[38] John G. Ratcliffe. *Foundations of hyperbolic manifolds*. Graduate Texts in Mathematics 149. Springer, 1994 (cit. on pp. v, 31, 41, 47, 74, 79, 80, 83).

[39] Walter Rudin. *Real and complex analysis*. McGraw-Hill, 1966.

[40] Jean-Pierre Serre. *Linear representations of finite groups*. Trans. by Leonard L. Scott. Graduate Texts in Mathematics 42. Springer-Verlag, 1977 (cit. on pp. 109, 111).

[41] Jean-Pierre Serre. *Trees*. Trans. by John Stillwell. Springer, 1980 (cit. on pp. 101, 102, 105, 107).

[42] Igor R. Shafarevich. *Basic algebraic geometry 1. Varieties in projective space*. 3rd ed. Springer, 2013 (cit. on pp. 109, 112).

[43] William P. Thurston. *The Geometry and Topology of Three-Manifolds*. Unpublished notes. 1979. URL: `http://library.msri.org/books/gt3m/` (cit. on pp. 61, 73, 79, 80).

[44] William P. Thurston. *Three-dimensional geometry and topology*. Ed. by Silvio Levy. Vol. 1. Princeton University Press, 1997 (cit. on pp. 74, 78).

[45] Vladlen Timorin. *Moore’s theorem*. 2010. arXiv: `1001.5140 [math.DS]` (cit. on p. 89).

Index