

INTRODUCTION TO HYPERBOLIC SURFACES

Notes of a course given in CIRM-Luminy, March 13-19, 2008

June 4, 2008

Chapter 1

Foreword

The aim of this small series of note is to give a concise and elementary introduction to hyperbolic surfaces. Starting from a synthetic point of view, we shall give the classification of compact oriented connected hyperbolic surfaces, define arithmetic surfaces introduce some elementary dynamical ideas and finally state some of the major conjectures in the subject.

The text is meant to be accessible to a student who knows about the cross ratio and the projective line, what is a covering space and a simply connected object.

This set of note is very preliminary and the result should evolved to a more correct version.

Many excellent textbooks or notes on the web exist and I would like in particular suggest to consult N. Bergeron remarkable course on a broader subject.

This text contains many errors, wrong statements, no bibliography and its orthograph and syntax are pathetic. So please do not make it circulate.

Contents

1	Foreword	3
2	Hyperbolic plane	7
2.1	Synthetic geometry	7
2.1.1	The complex projective line and how to build hyperbolic planes	8
2.1.2	The upper half space model	9
2.2	More geometric features: figures and distances	9
2.2.1	Triangles and ideal triangles	10
2.2.2	Right-angled hexagons	10
2.2.3	Regular polygons	10
2.3	The Riemannian interpretation	10
2.3.1	The length of a curve	10
2.3.2	Area	11
2.3.3	The Gauss-Bonnet formula	11
3	Hyperbolic surfaces	13
3.1	Surfaces	13
3.1.1	Hyperbolic surfaces	13
3.1.2	Local isometries	13
3.2	Construction of hyperbolic surfaces by gluing	14
3.2.1	Gluing two length metric spaces	14
3.2.2	Construction of closed surfaces	14
3.3	Cutting and classification	15
3.3.1	Simply connected surfaces	15
3.3.2	Curves on surfaces	16
3.3.3	Characterisation of pair of pants	16
3.3.4	Final statement	18
4	Discrete subgroups and closed surfaces	21
4.1	Monodromies of hyperbolic structures	21
4.1.1	Teichmüller space and space of representations	22
4.2	Arithmetic surfaces	22
4.2.1	Commensurator group	22
4.2.2	Arithmetic hyperbolic surfaces	23
4.2.3	Hecke correspondences and arithmetic dynamics	23

5	Dynamics	27
5.1	The action on the boundary at infinity	27
5.2	The unitary tangent bundle and flows	29
5.2.1	The Anosov property	30
5.3	Measures	31
5.3.1	Equidistribution and growth of geodesics	31
5.4	Harmonic functions	32
5.4.1	The trace formula	32
5.5	Conjectures	33

Chapter 2

Hyperbolic plane

2.1 Synthetic geometry

The complete geometry of the hyperbolic plane can be recovered synthetically from several features, namely *lines* and *boundary at infinity*. Let us define

Definition 2.1.1 [HYPERBOLIC PLANE] *A set A together with a*

- 1. a subset B called the boundary at infinity with a cyclic order,*
- 2. a family of lines which are subsets of A ,*

is a weak (completed) hyperbolic plane if

- 1. given any two distinct points in A there exists a unique line passing through two points of A ,*
- 2. Any line intersects the boundary at infinity in exactly two points.*
- 3. Two lines with four intertwined ends intersects in a point in $A \setminus B$.*
- 4. Two lines with four non intertwined ends do not intersect.*
- 5. If x is a point in B , and ι_x is the involution of B exchanging the end points of the lines through y , if x is distinct than y then the hyperbolic translation $\iota_x \circ \iota_y$ has exactly two fixed points in B : the end points of the line passing through x and y .*

The set A is completed hyperbolic plane if moreover the set B – together with its cyclic order – is equipped with a cross ratio that identifies it with $\mathbb{P}(P)$ for some real two plane P , and if furthermore this cross ratio is invariant under the involution above.

In this case, we use the following notation, we denoted the completed hyperbolic plane by $\overline{\mathbb{H}^2}$, its boundary at infinity by $\partial_\infty \mathbb{H}^2$, and the hyperbolic plane itself is $\mathbb{H}^2 = \overline{\mathbb{H}^2} \setminus \partial_\infty \mathbb{H}^2$.

EXERCISE : All hyperbolic planes are geometrically equivalent, by which we mean that there exists a bijection sending lines to lines, boundary at infinity to boundary at infinity and preserving the cross ratios. The idea is that any point in $A \setminus B$ defines an involution of the boundary

at infinity which preserves the cross ratio by exchanging the end points and which is different than the identity. Conversely any such involution can be produced that way. Therefore we have identified $A \setminus B$ with the set of involutions of the projective line preserving the cross ratio and lines can be now defined accordingly using the hyperbolic translations.

Then almost by definition the group $\text{PSL}(P)$ acts transitively on the hyperbolic plane preserving the geometry.

It remains to construct hyperbolic planes and we shall that they exists thanks to the existence of complex numbers. If Euclid would have known complex numbers, as well as understood Desargues, he would have construct hyperbolic geometry as in the next section ...

2.1.1 The complex projective line and how to build hyperbolic planes

Let E be a vector space of dimension 2 over a field $\mathbb{K}$. For a geometer the most important feature of the group $\text{PSL}(E, \mathbb{K})$ is its action on the projective line $\mathbb{P}(E)$, action which is characterised by the fact it preserves the cross ratio. In an affine chart, the action is given by homographies

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} x = \frac{ax + b}{cx + d}.$$

The cross ratio itself is

$$[x_1, x_2, x_3, x_4] = \frac{x_1 - x_2}{x_1 - x_4} \cdot \frac{x_3 - x_4}{x_3 - x_2}.$$

For $\mathbb{K} = \mathbb{C}$, the fact that $\mathbb{C}$ is an extension of $\mathbb{R}$ defines an extra geometry on the projective line $\mathbb{P}(E)$.

The starting observation is that every complex line in E is a real two plane in E seen as a real vector space.

Conversely, every real two plane P in E determines

1. a *circle* C_P in $\mathbb{P}^1(\mathbb{C})$ which is the set of complex lines that intersects P , and which is identified with $\mathbb{P}^1(\mathbb{R})$.
2. in the case, P is not a complex line, an *involution* preserving C . This involution σ_P reflect the Galois group of the extension of $\mathbb{R}$ by $\mathbb{C}$ and can be characterised in several ways. The more geometric one is the following. if P is a real two-plane and D is a complex line, $\sigma(D)$ is characterised by

$$\forall D_1, D_2, D_3, \in C_P, [D_1, D_2, D_3, D] = \overline{[D_1, D_2, D_3, \sigma(D)]}.$$

When $P = \mathbb{R}^2$ in $E = \mathbb{C}^2$, the corresponding involution is given by complex conjugation. All these quantities are equivariant under the action of $\text{PSL}(E, \mathbb{C})$.

The product of two such involutions preserves the cross ratio and is therefore a projective transformation. We finally define the *angle* between two intersecting and oriented circles C_P and C_Q as the argument of the trace of the product of the two involutions σ_P and σ_Q .

We can now proceed to the construction of the hyperbolic plane.

Proposition 2.1.2 *Let P be a real two-plane in a complex two-vector space – for instance $\mathbb{R}^2$ – itself. The completed hyperbolic plane is the set of complex lines L so that*

$$\forall D_1, D_2, D_3, \in C_P, \Im([D_1, D_2, D_3, D]) \geq 0.$$

A hyperbolic line or geodesic is the intersection of a circle orthogonal to P . The boundary at infinity is C_P .

2.1.2 The upper half space model

Using a chart, we end up with a more familiar picture: the hyperbolic plane is the upper half plane in $\mathbb{C}$.

- The *completed hyperbolic plane* is

$$\overline{\mathbb{H}^2} = \{z \in \mathbb{C}, \Im(z) \geq 0\} \cup \{\infty\}.$$

- The *hyperbolic plane* is

$$\mathbb{H}^2 = \{z \in \mathbb{C}, \Im(z) > 0\}.$$

- The *boundary at infinity* is

$$\partial_\infty \mathbb{H}^2 = \{z \in \mathbb{C}, \Im(z) = 0\} \cup \{\infty\}.$$

- The *lines* are either half circles orthogonal to the real axis or half lines orthogonal to the real axis completed by ∞ .

Then the natural action of $\mathrm{PSL}(2, \mathbb{R})$ is given by homographies

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} x = \frac{ax + b}{cx + d}.$$

Observe finally that the angles – in the upper half plane model – between two oriented geodesics can now be defined using the cross ratios of their end points. In particular, two geodesics intersect orthogonally if their end points form a *harmonic division*, that is if their cross ratio is -1.

2.2 More geometric features: figures and distances

If two geodesics D and L do not intersect, we define the *distance* between them as the logarithm of the cross ratio of their end points. We define the *distance* between two points as the distance between the geodesics that are orthogonal. It is not yet clear that this is a distance, we shall prove that later in some more economic way. For the moment we admit that this is indeed a distance. *Hyperbolic circles* – later on called simply circles – are the circles in our geometric model that do not intersect the boundary at infinity, *horocycles* are circles that intersect the boundary at infinity at exactly one point

A *half space* is a complementary region to a geodesic, and *wedge* is one of the complementary region of two oriented geodesic intersecting orthogonally.

A *convex polygon* is the intersection of half spaces. Among them are triangles, hexagons etc ... two points in a convex polygons are joined by a geodesic arc inside this polygon

2.2.1 Triangles and ideal triangles

An *ideal triangle* is a triangle with three points at infinity, a *2/3-ideal triangle* has two points at infinity and a *1/3-ideal triangle* has one. All ideal triangles are congruent.

Proposition 2.2.1 *Given any number a , b and c satisfying the triangles inequalities there exist a unique triangle – up to the action of $\mathrm{PSL}(2, \mathbb{R})$ – whose length are a , b and c .*

Idea: we fix a point and one segment of length a , we consider the segment at angle θ and length b then the distance between the end points is an increasing function of θ . A continuity argument shows this is a bijection onto $[|a - b|, a + b]$.

2.2.2 Right-angled hexagons

Proposition 2.2.2 *Given any number a , b and c , there exist a unique right angled hexagon – up to the action of $\mathrm{PSL}(2, \mathbb{R})$ – whose length of non intersecting edges are a , b and c .*

idea : same idea as above.

2.2.3 Regular polygons

Proposition 2.2.3 *Given any integer $n > 4$, there exist a unique – up to the action of $\mathrm{PSL}(2, \mathbb{R})$ – regular right-angled n -gon.*

idea : same idea as above.

2.3 The Riemannian interpretation

2.3.1 The length of a curve

We define the length of a parametrised curve $c = (x, y) : [a, b] \rightarrow \mathbb{H}^2$ in the hyperbolic plane in the upper half model as

$$\ell(c) = \int_a^b \frac{\sqrt{\dot{x}^2 + \dot{y}^2}}{y} dt.$$

Then, the following facts are true

- The length of the curves is invariant under the action $\mathrm{PSL}(2, \mathbb{R})$. *Hint: check the invariance under well chosen subgroups if you are lazy*
- Geodesics minimise the length of the curve. *Hint: consider first the case of a curve orthogonal to the real axis.*
- The distance between two points is

$$d(x, y) = \inf\{\ell(c) \mid c \text{ joins } x \text{ to } y\}.$$

similarly for distance between geodesics.

- The hyperbolic space is a *length space* in the sense that we can recover the length from the distance:

$$\ell(c) = \sup\left\{\sum_i d(c(x_i), c(x_{i+1})) \mid 0 = x_0 < x_1 < \dots < x_n = 1\right\}.$$

Finally, one recover the boundary at infinity from this picture. We say two oriented geodesics are *asymptotic* if given two arc lengths parametrisation of these geodesics $t \rightarrow \gamma_1(t)$ and $t \rightarrow \gamma_2(t)$ then

$$\lim_{t \rightarrow +\infty} \sup (d(\gamma_1(t), \gamma_2(t))) < \infty.$$

Then two oriented geodesics are asymptotic precisely if they have the same end point at $+\infty$.

2.3.2 Area

The area of a measurable set A is

$$\text{area}(A) = \int_A \frac{1}{y^2} dx dy.$$

2.3.3 The Gauss-Bonnet formula

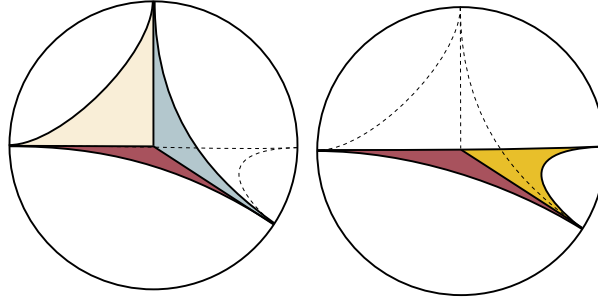


Figure 2.1: Gauss additivity

Theorem 2.3.1 *The area of a triangle is its angle defect that is π minus the sum of the interior angles of the triangle.*

PROOF : We shall follow Gauss approach. We first check by a direct computation that the area of an ideal triangle is π . Then let $A(\theta)$ the area of a $2/3$ -ideal triangle with angle $\pi - \theta$. Gauss observation is $A(\theta)$ is an additive function hence a multiple of θ , hence θ by the normalisation. This is done geometrically as follows in Figure 2.1: the area of the union of the pink and blue triangles is the area of yellow triangle, since they are both π minus the area of the purple triangle. Once this is done the rest follows.Q.E.D.

As an exercise, define an area just using the angle defect and check which properties the angle defect should have and prove them.

As a special we also see that the area of a right angle hexagons is 2π and that of of regular right-angled n -gon is $\frac{\pi}{2}n$.

Chapter 3

Hyperbolic surfaces

3.1 Surfaces

3.1.1 Hyperbolic surfaces

A *hyperbolic surface* is a complete metric space M such that every point in M has a neighbourhood isometric to an open set of the hyperbolic plane.

A *hyperbolic surface with totally geodesic boundary* is a complete metric space M such that every point in M has a neighbourhood isometric to an open set of the hyperbolic plane, or an hyperbolic half plane.

A *hyperbolic surface with totally geodesic boundary and right angles* is a complete metric space M such that every point in M has a neighbourhood isometric to an open set of the hyperbolic plane, or an hyperbolic half plane, or a right-angled wedge.

It follows from the definitions that hyperbolic surfaces are length spaces. This has an interesting consequence:

Lemma 3.1.1 *Two points in a hyperbolic surface with totally geodesic boundary and right angles can be joined by a geodesic whose length realise the distance between the two points.*

PROOF : use the fact that hyperbolic surfaces are locally compact spaces. Then produce a length minimising sequence of curves parametrised by arc length and prove they converge to a geodesic by using Arzela-Ascoli lemma. Q.E.D.

3.1.2 Local isometries

We begin with the following two lemmas

Lemma 3.1.2 *Every map from a subset U of $\mathbb{H}^2$ containing more than three points to $\mathbb{H}^2$ preserving the distance is the restriction of an isometry. Every orientation preserving map from an open connected subset of $\mathbb{H}^2$ to $\mathbb{H}^2$ preserving the length of curves is the restriction of a homography.*

PROOF : Prove the first easy statement then the second one.Q.E.D.

Lemma 3.1.3 *Let ϕ be a local isometry from S to Σ which both are hyperbolic surfaces, then ϕ is a covering.*

PROOF : take a ball in S isometric to a ball in $\mathbb{H}^2$. Then, use Lemma 3.1.1 to show that the preimage of this ball B is a union of disjoint balls all isometric to B . Q.E.D.

3.2 Construction of hyperbolic surfaces by gluing

3.2.1 Gluing two length metric spaces

We can glue two length spaces provided we have a gluing map that preserves the distances, by defining the length of any curves as the sum of the length in each part.

Then one checks that the gluing two hyperbolic half-spaces along their boundary leads to the hyperbolic plane, and gluing two right-angle wedges leads to the hyperbolic half plane.

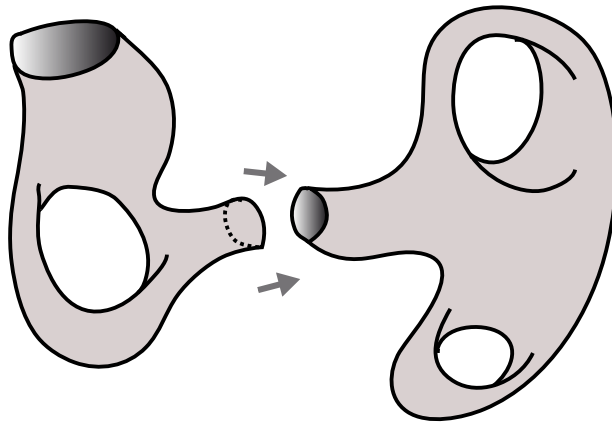


Figure 3.1: Gluing surfaces

3.2.2 Construction of closed surfaces

We can therefore construct hyperbolic surfaces of area $4\pi n$ using $3n$ -positive real parameters, and $3n$ -angles – i.e elements of $\mathbb{R}/\mathbb{Z}$. Moreover, the surface is given together with an extra topological structure namely an *decomposition into pair of pants*. The construction run as follows. First, we construct pairs of right-angled hexagons, fixing the boundary length. Then, gluing two hexagons together we obtain a hyperbolic pair of pants whose boundary length are prescribed. Then we glue these pair of pants together using a prescribed identification of the boundaries. We can isometrically identify each boundary component of length a with

$\mathbb{R}/a\mathbb{Z}$, in a canonical way: sending a chosen corner of the previous hexagon to zero. Then the gluing between two boundary components is determined by one parameter in $\mathbb{R}/a\mathbb{Z}$.

Two questions remain

1. Are these surfaces isometric ?
2. Are all compact surfaces obtained this way ?

3.3 Cutting and classification

Let us state our main result, which answers partly the question above.

Theorem 3.3.1 *Every compact oriented connected hyperbolic surface is of area $4\pi n$ and can be obtained be decomposed into $2n$ hexagons glued as pair of pants. Moreover, this decomposition, as well as the $3n$ gluing parameters, is fixed as soon as we fix $3n$ homotopy classes of pairwise non intersecting simple curves on S .*

3.3.1 Simply connected surfaces

Theorem 3.3.2 *The universal cover of a hyperbolic surface with geodesic boundary and right-angles is isometric to a convex non compact polygon in the hyperbolic plane.*

We state a useful corollary.

Corollary 3.3.3 *Let S be a hyperbolic surface with right-angled. Then any loop based at a point is homotopic to a geodesic loop. Conversely, any curve homotopic to a geodesic loop, is not homotopic to zero.*

PROOF : We sketch the proof here. By considering the double of the hyperbolic surface, it suffices to prove the result for hyperbolic surface without boundary which we do now. Let $\{U_i\}_{i \in I}$ be a cover of a hyperbolic surface so that every U_i is isometric to a ball in the hyperbolic plane.

Fix an index i_0 , a *path* starting from i_0 is a finite sequence of indices $\{i_0, i_1, \dots, i_n\}$ so that U_{i_j} intersects $U_{i_{j+1}}$. We fix once and for all an isometry f_0 of U_{i_0} to a ball in the hyperbolic plane.

The first remark, which follows from Lemma 3.1.2 is the following: given a path P there exists a unique family of isometries $f_{i_k}^P$ of U_{i_k} in $\mathbb{H}^2$ so that $f_{i_k}^P$ coincide with $f_{i_{k-1}}^P$ on the intersection. We call $f_{i_n}^P$ the *final isometry*

Then, we say two paths with fixed extremity i_n are *homotopic* if they can be obtained from each other after a succession of the following elementary operation: *deletion* which is just to remove an index (whenever it is possible) or *insertion* which is the converse operation. Using Theorem 3.3.2, it is easy to see that two homotopic path defines the same final isometries.

Now we can proceed in two ways

1. either use our little knowledge of homotopy theory and notice that in a simply connected hyperbolic surface, each "discrete" path is homotopic to any another path,
2. or construct directly a hyperbolic surface Σ with a covering $\{V_j\}_{j \in J}$ together with a map f from Σ to S so that

- (a) all paths in Σ are homotopic to zero,
- (b) for any $j \in J$, there exists $i \in I$, so that f is an isometry from V_j to U_i .

This is a fairly easy construction obtained by gluing. Then, by Lemma 3.1.3 Σ is S .

It follows that we can extend uniquely the map f_{i_0} to a map f from S to $\mathbb{H}^2$. Then by Lemma 3.1.3 again, f is a covering, hence an isometry.

Q.E.D.

3.3.2 Curves on surfaces

We summarise here the property of curves on compact hyperbolic surfaces with boundary that we prove.

Theorem 3.3.4 *1. Given two point p and q in S and a path c from p to q there exists a unique geodesic joining p to q . Moreover this geodesic minimises the length of all curves joining p to q .*

2. Given a closed curve c , there exists a unique closed geodesic γ_c freely homotopic to c , whose length furthermore minimise the length of all curves freely homotopic to c . Moreover

- (a) *if c is simple so is γ_c ,*
- (b) *if c_1 and c_2 do not intersect and if furthermore no non trivial powers of these curves are freely homotopic, γ_{c_1} and γ_{c_2} do not intersect.*

PROOF : The first statement is a consequence of our description of the universal cover.

For the second statement, we first find for every x in S , geodesics loops freely homotopic to c . Then we can find using the compactness of S a loop of minimal length, and this loop has to be a closed geodesics.

For the last statement we first make the following important remark: *if a curve stays at a finite distance of a geodesic, then it converges at infinity to the end points of geodesics in the upper half-space model.* Therefore, if a curve is freely homotopic to a geodesic loop, any of its lift in the universal cover stay at finite distance to a geodesic and hence converges at infinity. Therefore, if the geodesics loops associated to the two curves intersects then we can find lifts of these geodesics that intersects, then their endpoint are intertwined, hence the corresponding lifts of the curves intersect too by Jordan theorem, and hence the two initial curves also intersect..

The statement about simple curves is an elaboration of this idea. Q.E.D.

3.3.3 Characterisation of pair of pants

We start with the following result that follows from the classification of surfaces.

Theorem 3.3.5 *Let S be a surface that is not topologically a pair of pants, then it contains a simple curve not freely homotopic to a boundary component.*

This result follows from the classification of surfaces but we sketch an independent proof. Let us start with a definition. Let $H^1(S) = \pi_1(S)/[\pi_1(S), \pi_1(S)]$. We say two closed curves are *homologous* if they coincide in $H^1(S)$. We observe that if curves are not cohomologous then, they are not freely homotopic/

STEP 1 As a corollary of the description of the universal cover of a hyperbolic surface in Theorem 3.3.2, we have

Proposition 3.3.6 *The fundamental group of a surface with at least one boundary component is free with at least two generators. Moreover if there are n boundary components, then any $n - 1$ of them are independent.*

PROOF : consider a polygonal fundamental domain, it is a convex set, the image of its boundary in the quotient is a graph, whose complementary region is simply connected and isometric to a convex polygon. This convex polygons as at least one side which lies in the boundary component, we can therefore retract this side to the other sides. The result follows. Q.E.D.

STEP 2

Proposition 3.3.7 *Let S be a compact hyperbolic surface with at most two totally geodesic boundary components. Then there exists a non trivial embedded closed geodesic different than the boundary.*

PROOF : We first need to master a technique: how to produce from a simple curve not homologous to a power of boundary components from a curve not homologous to boundary components. The idea is as follows : you find a curve with transverse self intersections – using broken geodesic arcs for instance – then you open up intersections. Then at least one of the simple curves obtained is not homologous to a power the boundary component, hence it is not freely homotopic to a power of boundary component. From the first step one easily produce on a surface with at most two boundary components, a curve not homologous to a boundary component

Q.E.D.

STEP 3

Proposition 3.3.8 *Let S be a compact hyperbolic surface with at least four totally geodesic boundary components. Then there exists a simple closed curve not freely homotopic to a power of a boundary.*

PROOF : Take an arc joining two boundary component and consider thicken it to obtain a simple curve that is homologous to the sum of the two boundary components. This curve is not homologous to any power of a given boundary component, hence not freely homotopic either. Q.E.D.

STEP 4 This last step will useful later on.

Proposition 3.3.9 *A surface S with three totally geodesic boundaries. Assume that S does not contain a simple curve not homotopic to any power of a boundary component. Then S can be decomposed uniquely into a pair of right-angled hexagons*

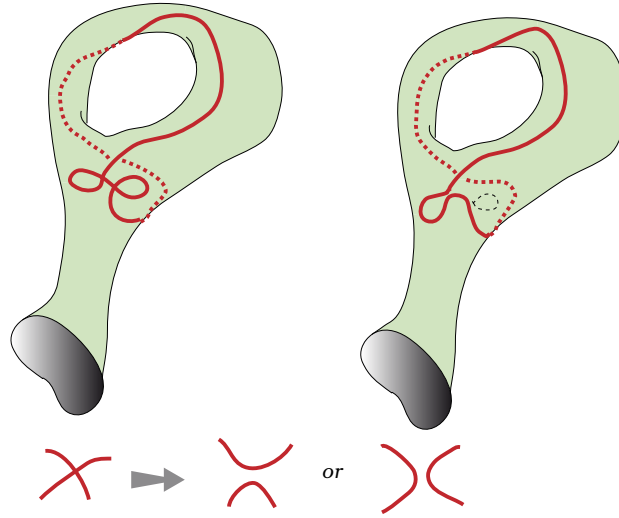


Figure 3.2: Opening up intersections

PROOF : We first build simple arcs joining the three components of the boundary. To do so, you find three arcs – for instance using broken geodesic – joining the three boundary components. with transverse intersection and you open up all intersections wisely that is proceeding recursively from one boundary component, and always choosing one of the opening up that does not go back to your original boundary. You delete loops and you end up with three arcs. Then you straighten them up using geodesics, that is you find the shortest arc homotopic to them (Hint: use a doubling of the surface). You observe that, by symmetry the geodesics intersect the boundary orthogonally. Then the complementary region is not connected.

We now prove that the remaining two connected components are simply connected. Let c be a loop in one of this region not homotopic to zero. We can assume after opening up intersections, that it is a simple loop, and after straightening up to an embedded closed geodesic, which is different than any of the boundary components since it stay in that complementary region. This contradict the hypotheses.

3.3.4 Final statement

As an exercise, prove Theorem 3.3.1 by cutting successively using simple closed curves. The process stops at a finite time and the remaining pieces are hyperbolic pair of pants. The stop occurs thanks to the following lemma

Lemma 3.3.10 *A surface with totally geodesic boundary has area which is a multiple of π .*

PROOF : we use the description of the universal cover and cut the surface into hyperbolic triangles. Then the sum of angle defects is the area, summing these angle defects of vertices show that the sum is a multiple of π . Q.E.D.

As another consequence, we obtain

Theorem 3.3.11 [GAUSS-BONNET FORMULA (BIS)] *Let S be a compact surface admitting a hyperbolic structure. Then two pair of pants decomposition have the same number of pair of pants and this number is even. If $\chi(S)$ is this opposite of this number called the Euler characteristic of S , then any hyperbolic structure on the surface has area $-2\pi\chi(S)$.*

Chapter 4

Discrete subgroups and closed surfaces

We now assume that S is oriented.

4.1 Monodromies of hyperbolic structures

Every surface gives rise to an embedding of the fundamental group of $\pi_1(S)$ with discrete image, moreover this group has no torsion. Indeed, every torsion element of $\mathrm{PSL}(2, \mathbb{R})$ fixes a point in the hyperbolic plane.

Conversely every torsion free subgroup of $\mathrm{PSL}(2, \mathbb{R})$ acts properly freely on the hyperbolic plane and is the monodromy of a – non necessarily compact – hyperbolic surface. In order to complete the picture, we recall

Lemma 4.1.1 [SELBERG] *Every finitely generated linear group possesses a finite index subgroup without torsion.*

It follows that every faithful representation of the fundamental group of a surface with discrete image is the monodromy of a hyperbolic structure. A little extra work shows that if moreover S is compact then $\mathbb{H}^2/\rho(\pi_1(S))$ is homeomorphic to S , so therefore

Proposition 4.1.2 *Every faithful representation of the fundamental group of a compact surface S with discrete image is the monodromy of a hyperbolic structure on S .*

EXERCISE: show that this last statement fails for a non compact surface.

Proposition 4.1.3 [BOREL DENSITY THEOREM] *Every monodromy of a finite volume hyperbolic surface is Zariski dense.*

PROOF : We prove it for a compact surface $S = \mathbb{H}^2/\Gamma$. From the description of hyperbolic surfaces, it follows that we can find two elements in Γ which generate different hyperbolic translations. Then, considering the very short list of algebraic subgroups of $\mathrm{PSL}(2, \mathbb{R})$, we see that this cannot happen unless the algebraic group is $\mathrm{PSL}(2, \mathbb{R})$ itself Q.E.D.

4.1.1 Teichmüller space and space of representations

The *Teichmüller space* $\tau(S)$ is the space of all representations of $\pi_1(S)$ with discrete image up to conjugacy. It is diffeomorphic to a ball, we almost proved it Let us introduce an invariant of representation of $\pi_1(S)$. For that, let us choose a presentation of $\pi_1(S)$

$$\pi_1(S) = \langle a_1, b_1, \dots, a_g, b_g \mid \prod_{i=1}^g g[a_i, b_i] = 1 \rangle,$$

where $[c, d] = cdc^{-1}d^{-1}$.

Observe now that $\mathrm{PSL}(2, \mathbb{R})/\mathrm{SO}(2) = \mathbb{H}^2$, hence that $\mathrm{PSL}(2, \mathbb{R})$ has the homotopy type of S^1 . We therefore have an exact sequence

$$\mathbb{Z} \rightarrow \widetilde{\mathrm{PSL}(2, \mathbb{R})} \rightarrow \mathrm{PSL}(2, \mathbb{R}) \rightarrow 0.$$

Let us choose a map σ from $\mathrm{PSL}(2, R)$ to $\widetilde{\mathrm{PSL}(2, \mathbb{R})}$ that splits that sequence, σ will actually never be continuous, nor a group morphism. Then we have

Proposition 4.1.4 [EULER CLASS] *Let ρ be a representation of $\pi_1(S)$ to $\mathrm{PSL}(2, \mathbb{R})$. The element*

$$e(\rho) = \prod_{i=1}^g g[\sigma(\rho(a_i)), \sigma(\rho(b_i))],$$

is an element of the centre of $\widetilde{\mathrm{PSL}(2, R)}$ which we identify to $\mathbb{Z}$. This number is independent of the choice of σ , of the presentation of $\pi_1(S)$ and is constant under local deformations of ρ . The number $e(\rho)$ is called the Euler class of the representation.

Let now $\chi(S)$ be the Euler characteristics of S . Then

Theorem 4.1.5 [MILNOR-WOOD INEQUALITY] *Let ρ be a representation of $\pi_1(S)$ to $\mathrm{PSL}(2, \mathbb{R})$. Then*

$$|e(\rho)| \leq |\chi(S)|.$$

We can use the Euler class to distinguish connected components of space of representations. More precisely

Theorem 4.1.6 [GOLDMAN] *The map from the space of connected components of $\mathrm{Hom}(\pi_1(S), \mathrm{PSL}(2, \mathbb{R}))$ to $\{\chi(S), \chi(S) + 1, \dots, -\chi(S)\}$ is a bijection. Moreover, monodromies of hyperbolic structures are exactly representations such that $|e(\rho)| = |\chi(S)|$.*

It follows from this theorem that we can check whether a representation is the monodromy of a hyperbolic structure just from a presentation of the group.

4.2 Arithmetic surfaces

4.2.1 Commensurator group

Let S be a hyperbolic surface $S = \mathbb{H}^2/\Gamma$. The commensurator group of a subgroup of Γ is

$$\mathrm{Comm}(\Gamma) = \{g \in \mathrm{PSL}(2, \mathbb{R}) \mid \#(g\Gamma g^{-1} \cap \Gamma) < \infty\}.$$

We check for instance that

$$\mathrm{Comm}(\mathrm{PSL}(2, \mathbb{Z})) = \mathrm{PSL}(2, \mathbb{Q}).$$

4.2.2 Arithmetic hyperbolic surfaces

By a non standard definition that rests on a deep result by Margulis, we define an *arithmetic hyperbolic surface* to be a finite volume hyperbolic surface whose commensurator of the fundamental group is dense in $\mathrm{PSL}(2, \mathbb{R})$.

Arithmetic hyperbolic surfaces are classified, they are more or less obtained through the following procedure: first we find a representation of $\mathrm{PSL}(2, \mathbb{Q})$ in $\mathrm{PSL}(n, \mathbb{Q})$ then Γ is the intersection of the image of $\mathrm{PSL}(2, \mathbb{Q})$ with $\mathrm{PSL}(n, \mathbb{Z})$. It is not straightforward to construct compact arithmetic surfaces but they exist. Moreover, by a simple cardinality argument, non arithmetic surfaces also exist.

Here is another remark.

Lemma 4.2.1 *If a surface is not arithmetic, then Γ has finite index in $\mathrm{Comm}(\Gamma)$.*

PROOF : Let H the closure of $\mathrm{Comm}(\Gamma)$. We first prove that it is discrete: indeed otherwise, every element in $\mathrm{Comm}(\Gamma)$ would fix the Lie algebra of the connected component of H of the origin. But this defines a Zariski closed condition. By Borel density theorem, $H = \mathrm{PSL}(2, \mathbb{R})$ hence S is arithmetic which is a contradiction. It follows that $\mathrm{Comm}(\Gamma)$ is discrete. Then we have a covering map from $\mathrm{PSL}(2, \mathbb{R})/\Gamma$, which is compact, to $\mathrm{PSL}(2, \mathbb{R})/\mathrm{Comm}(\Gamma)$. Hence, the fibres of this map are finite sets which exactly means that Γ has finite index in $\mathrm{Comm}(\Gamma)$. Q.E.D.

4.2.3 Hecke correspondences and arithmetic dynamics

The main feature of arithmetic surfaces are the existence of many correspondences. A *finite correspondence* between two sets X and Y is a subset of Z in the product $X \times Y$ so that the preimage of very point in each of the factor is finite and non empty. In particular, an element g in the commensurator group of a hyperbolic surface S gives rise to such a correspondence which is furthermore a local isometry see Figure 4.1. Indeed, we consider the map of $\mathbb{H}^2$ into $S \times S$ given by $x \rightarrow (\pi(x), \pi(gx))$ where π is the covering map. To say g is in the commensurator group, is just to say that this map is a covering map of compact image and that its image is a correspondence Z_g . Actually, the correspondence is just determined by the class of g in $\Gamma \backslash \mathrm{Comm}(\Gamma)/\Gamma$.

We now define *correspondences for hyperbolic surfaces* to be correspondences which are local isometries. The dichotomy between arithmetic surfaces and non arithmetic surfaces is then the dichotomy between finitely many and infinitely many self correspondences.

A self correspondence gives rise to two types of dynamics. First *quantum dynamics* acting on the space of L^2 functions on S . So if p_1 and p_2 are the two projections – of degree q – of the correspondence $Z \subset S \times S$ on each factor, then

$$\mathcal{H}_g(f)(x) = \frac{1}{q} \sum_{z \in p_1(x)} f(p_2(z)),$$

is a self adjoint operator called the *Hecke operator* of the correspondence.

Secondly, we can associate *classical dynamics*. There is a classical dynamical way to turn non bijective map or more generally correspondence into a bijective map. So, to settle notation,

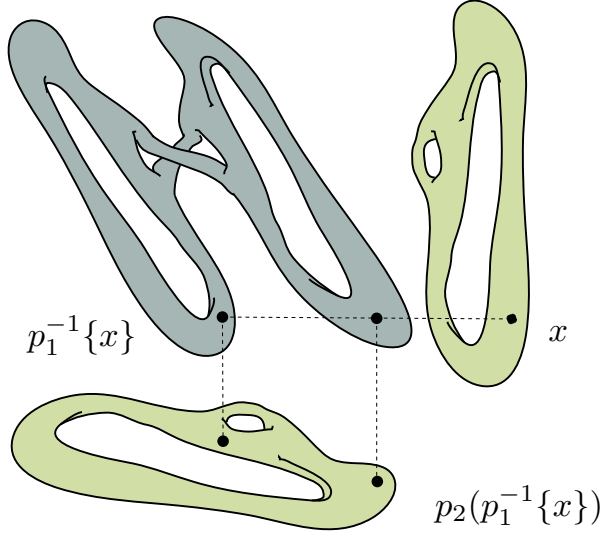


Figure 4.1: A correspondence

let Z be a correspondence and p_1 and p_2 be the two projections, we say $x\mathcal{R}y$ if $p_1^{-1}(x)$ intersects $p_2^{-1}(y)$. Then we consider the set

$$\mathcal{L}_Z = \{f : \mathbb{Z} \rightarrow S \mid f(n)\mathcal{R}f(n+1)\} \subset S^{\mathbb{Z}}.$$

The *shift* σ is the map from $\mathcal{L}_Z$ to itself given by $\sigma(f)(n) = f(n+1)$. The shift is now a homeomorphism and its dynamics reflect that of the correspondence.

This construction is not sufficient for our purpose. We indeed would like to see all correspondences as acting in the same space. Let us first take a look at the space $\mathcal{L}_Z$ in special case.

Let $P_{[n,p]}$ be the map from $\mathcal{L}_Z$ to S^{n-p} given by

$$f \rightarrow (f(n), \dots, f(p)).$$

By construction the image of this map is a compact surface, which we call $S_{[n,p]}$, in the product and moreover any restriction map is a covering map. We can therefore describe $\mathcal{L}_Z$ as a limit of some coverings.

We generalise this construction. Let S be a compact hyperbolic surface. The *Sullivan solenoid* $\mathcal{S}(S)$ is the "limit" of all coverings of S . Let us give a definition as a set. The *Sullivan solenoid* is the set of sequences $\{(x_n, S_n, p_n)\}_{n \in \mathbb{N}^*}$ so that x_n is a point in S_n , p_n is a covering from S_n to S_{n-1} – where by convention $S_0 = S$ – such that $p_n(x_n) = x_{n-1}$ up to the following *solenoid equivalence*: two sequences $\{(x_n^0, S_n^0, p_n^0)\}_{n \in \mathbb{N}}$ and $\{(x_n^1, S_n^1, p_n^1)\}_{n \in \mathbb{N}}$ are equivalent if there exists a third one $\{(y_n, \Sigma_n, q_n)\}_{n \in \mathbb{N}}$ together with covering maps q_n^i from Σ_n to S_n^i satisfying the commuting diagram conditions

$$\begin{aligned} q^i(n)(y_n) &= x_n^i, \\ p_n^i \circ q_{n+1}^i &= q_n^i \circ p_{n+1}^i. \end{aligned}$$

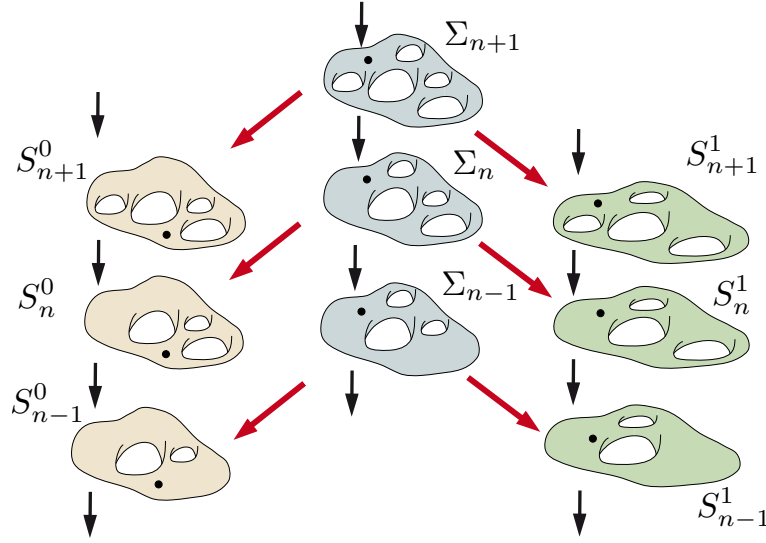


Figure 4.2: Solenoid equivalence

We give two alternate descriptions that describe the topology of this set

1. The combinatorial construction runs as follows. Let W_S be the 2-dimensional complex, whose vertices are surfaces which are finite covers of S , oriented edges correspond to covering between the extremities, and faces correspond to commuting diagrams of coverings. Let Z_S be the universal cover of this complex, and V_S be the set of vertices of this graph, which we consider as surfaces. If e is an edge of V_S from e^- to e^+ , then it gives rise to a covering p_e from e^- to e^+ seen as surfaces. Then

$$\mathcal{S}(S) = \{(x_\Sigma)_{i \in V_S} \mid x_\Sigma \in \Sigma, \ p_e(x_{e^+}) = x_{e^-}\}.$$

2. Alternatively, we can describe the Sullivan solenoid as fibre bundle over S whose structure group is the pro-finite completion of the fundamental group of S . The Sullivan solenoid is the "universal cover for finite covers", in the sense that it solves a universal problem for finite covers of S .

As a topological space, the *Sullivan solenoid* $\mathcal{S}(S)$ is a *hyperbolic laminated space* as in Figure 4.3: every point has a neighbourhood – called chart – which is homeomorphic to a product a ball in $\mathbb{H}^2$ with a topological space, such that moreover the coordinates changes when we change charts are isometries on the hyperbolic factor. A *leafwise isometry* of a hyperbolic laminated space is a homeomorphism that is a local isometry on the hyperbolic factors.

Now, here is the fundamental though obvious remark.

Lemma 4.2.2 *Let p a local isometry from S to Σ . Then the corresponding injection of $\mathcal{S}(S)$ to $\mathcal{S}(\Sigma)$ is a leafwise isometry.*

PROOF : use the fact that you can induce coverings. Q.E.D.

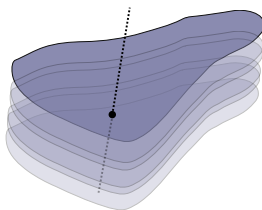


Figure 4.3: A small open set in a laminated space

As a consequence, a correspondence on a surface acts as a leafwise isometry on the Sullivan solenoid. Therefore, the Sullivan solenoid of an arithmetic surface has very rich dynamics.

As an exercise, we ask the reader to interpret the universal solenoid for a finite index torsion free subgroup of $\mathrm{PSL}(2, \mathbb{Z})$ in adélic terms.

Chapter 5

Dynamics

In this chapter, Γ will be a discrete subgroup of $\mathrm{PSL}(2, \mathbb{R})$ so that $\mathbb{H}^2/\Gamma$ is a compact hyperbolic surface.

5.1 The action on the boundary at infinity

We begin by studying the action of Γ on the boundary at infinity $\partial_\infty \mathbb{H}^2$ of $\mathbb{H}^2$. Every element of γ corresponds to a closed geodesic and will therefore preserve exactly two points at infinity $\{\gamma^-, \gamma^+\}$.

Moreover, every γ in Γ has *north-south dynamics* meaning that the sequence of iterates of any point x in $\partial_\infty \mathbb{H}^2$ different than γ^- converges to γ^+ , as in Figure 5.1.

$$\lim_{n \rightarrow \infty} (\gamma^n(x)) = \gamma^+.$$

We now use the north-south dynamics to show three crucial properties of the action of Γ on the boundary at infinity.

Lemma 5.1.1 [MINIMALITY] *Every orbit of Γ is dense on $\partial_\infty \mathbb{H}^2$.*

PROOF : Let F be a closed Γ invariant set in $\partial_\infty \mathbb{H}^2$ and E be the *convex envelope* of F in $\mathbb{H}^2$, that is the union of all geodesics joining points in F . The set E is a closed convex polygon which is Γ invariant. Let d be the function on $\mathbb{H}^2$ defined as a distance to E . Then d is Γ invariant. However d is unbounded as we see from taking a geodesic orthogonal to one of the boundary component of the polygon. This contradicts the compactness of $\mathbb{H}^2/\Gamma$. Q.E.D.

Lemma 5.1.2 [DENSITY OF END POINTS] *The set of end points of geodesics $\{(\gamma^+, \gamma^-) \mid \gamma \in \Gamma\}$, is dense in $\partial_\infty \mathbb{H}^2 \times \partial_\infty \mathbb{H}^2$.*

PROOF : By the previous lemma the set $\{\gamma^+ \mid \gamma \in \Gamma\}$ is dense in $\partial_\infty \mathbb{H}^2$. Let now (x, y) be pair of points in $\partial_\infty \mathbb{H}^2 \times \partial_\infty \mathbb{H}^2$, we can therefore find a pair of distinct points (η^-, γ^+) associated to elements η and γ . We remark that if two elements α and β of the group are such that $\alpha^+ = \beta^+$ then $\alpha^- = \beta^-$ (*Hint*: use the compactness of $\mathbb{H}^2/\Gamma$).

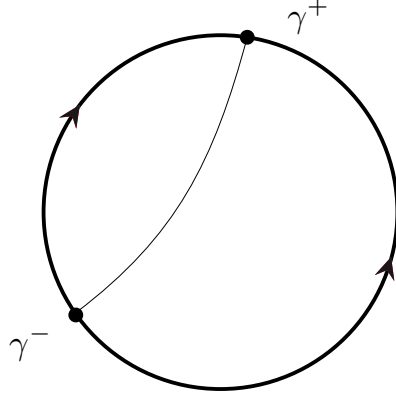


Figure 5.1: North-south dynamics

We therefore assume that all points $\eta^\pm, \gamma^\pm$ are distinct. The final remark is that

$$\lim_{n \rightarrow \infty} (\gamma^n \eta^n)^+ = \gamma^+,$$

and symmetrically

$$\lim_{n \rightarrow \infty} (\gamma^n \eta^n)^- = \gamma^-,$$

The process is describe in Figure 5.2 Let U be a small neighbourhood of γ^+ . Since γ^+ is different than η^- a high power of η will send U to a very small neighbourhood V of η^+ . Since η^+ is different than γ^- a high power of γ will send V to a even smaller neighbourhood of γ^+ . It follows that $\xi_n = \gamma^n \eta^n$ maps U into itself. Therefore it has a fixed point in U . This point is necessarily the attractive fixed point of ξ_n . This is what we wanted to prove.

Q.E.D.

Here is another consequence of this north-south dynamics.

Lemma 5.1.3 [HYPERBOLIC STABILITY] *Let S be a compact surface. Let ρ_1 and ρ_2 be two representations of $\pi_1(S)$ in $\text{PSL}(2, \mathbb{R})$ which are monodromies of hyperbolic structures on S . Then the two corresponding actions on $\partial_\infty \mathbb{H}^2$ are conjugate. More precisely there exists a unique – non smooth – homeomorphism Φ of $\partial_\infty \mathbb{H}^2$ so that*

$$\forall x \in \partial_\infty \mathbb{H}^2, \forall \gamma \in \pi_1(S), \Phi \circ \rho_1(\gamma)(x) = \rho_2(\gamma) \circ \Phi(x).$$

PROOF : let E_1 be E_2 be the set of end points of closed geodesics in ∂_∞ of respectively $\rho_1(\pi_1(S))$ and $\rho_2(\pi_1(S))$. Then we have a map Φ - *a priori* not continuous – from E_1 to E_2 that sends $\rho_1(\gamma)^+$ to $\rho_2(\gamma)^+$ (please check that Φ is well defined). Then Φ preserves the

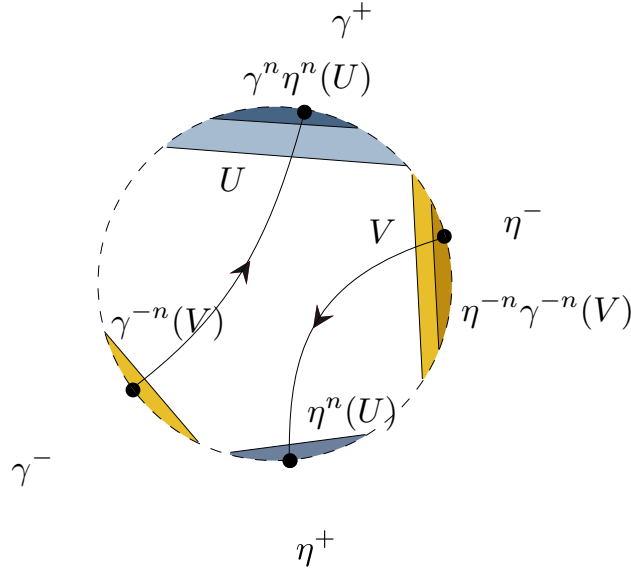


Figure 5.2: Density of pair of fixed points

intertwining of points, since intersecting geodesics correspond to intersecting geodesics. It follows by density of E_1 and E_2 that Φ extends to a continuous map. The conjugacy follows from these considerations. Q.E.D.

5.2 The unitary tangent bundle and flows

The group $\mathrm{PSL}(2, \mathbb{R})$ is the group of orientation preserving isometries of the hyperbolic plane. Moreover the group of orientation preserving isometries acts faithfully and transitively on the unitary tangent bundle US which we define equivalently as the set U of *pairs* (x, L) , where x is a point on the oriented geodesic L , or the set $\mathcal{T}$ of oriented distinct triples (x, y, z) of points in the boundary at infinity.

The *geodesic flow* is the one parameter subgroup $\{\phi_t\}$ of homeomorphisms of US , that sends (x, L) to (y, L) so that $[L^-, x, L^+, y] = t$. The *stable horocycle flow* is the one parameter subgroup that sends (x, L) to (u, M) so that L and M has a common positive end point z , and x and u are joined by the horocycle centred at z of length one. The following commutation rule, see Figure 5.3 holds and identify these flows with 1-parameter groups of matrices.

$$\phi_t \circ H_s \circ \phi_{-t} = H_{e^t s}. \quad (5.1)$$

Hence, the geodesic flows and horocycle flows are just the actions of the one-parameter diagonal and upper triangle groups of matrices on $\mathrm{PSL}(2, \mathbb{R})$ respectively.

Closed geodesics can – and will – now be interpreted as closed orbits of the geodesic flow.

EXERCISE Write the action of these flows on $\mathcal{T}$ using only the cross ratio.

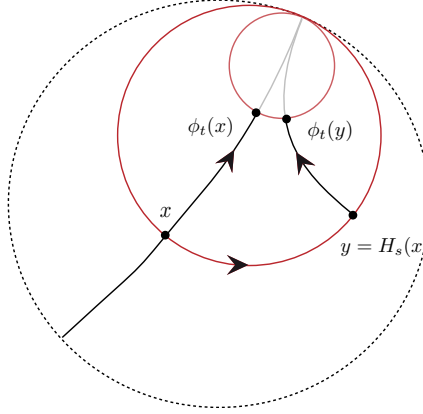


Figure 5.3: Commutation rule

5.2.1 The Anosov property

Let US be the unitary tangent bundle of the surface S , which from the discussion above is a quotient of $\mathrm{PSL}(2, \mathbb{R})$ by a discrete group Γ . We therefore have three flows on US .

1. The geodesic flow ϕ_t
2. The stable horocycle flow whose orbits we call *stable leaves*.
3. The unstable horocycle flow whose whose orbits we call *unstable leaves*, obtained by interchanging the role of end points.

Then the commutation rules (5.1) translate into the *Anosov property* of the geodesic flow, which we try to depict in Figure 5.4

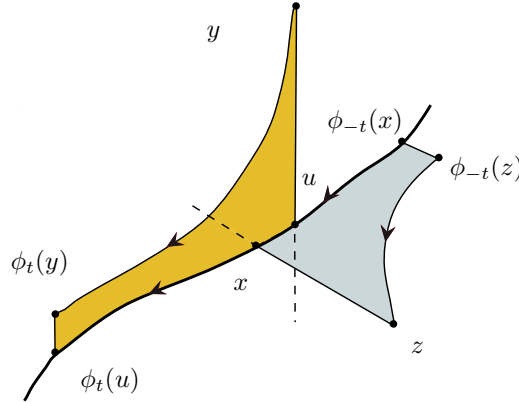


Figure 5.4: Anosov property

:

1. Two points on the same stable leaf get closer under a *positive* action of the geodesic flow.
2. Two points on the same unstable leaf get closer under a *negative* action of the geodesic flow.

This property is the translation for the geodesic flow of the north-south dynamics of the action of the fundamental group.

The Anosov property has a crucial consequence. We say a sequence of points $\{x_n\}_{n \in \mathbb{N}}$ in US is an ϵ -pseudo-orbit if there exists a positive number T so that for all n , the distance $d(x_n \phi_T(x_{n-1})) \leq \epsilon$.

Lemma 5.2.1 [SHADOWING LEMMA] *For any α , there exists ϵ -small enough, so that every ϵ -pseudo orbit is α close to an orbit*

The shadowing lemma implies the density of the reunion of all closed orbits which is also a consequence of Lemma 5.1.2.

5.3 Measures

5.3.1 Equidistribution and growth of geodesics

The unitary tangent bundle has a probability measure μ_0 – that we call the *Lebesgue* measure – which comes from the Haar measure of $\mathrm{PSL}(2, \mathbb{R})$ and which is invariant under the geodesic flow. Every closed geodesic γ also defines a geodesic flow invariant probability measure μ_γ on US by the formula

$$\int f d\mu_\gamma = \frac{1}{\ell(\gamma)} \int_0^{\ell(\gamma)} f(\phi_t(x)) dt,$$

where $\ell(\gamma)$ is the length of γ and x is any point of γ .

These measures are related by the following deep result

Theorem 5.3.1 [BOWEN, MARGULIS] *The closed geodesics are equidistributed with respect to the Lebesgue measure*

$$\lim_{T \rightarrow \infty} \frac{1}{\#(\Gamma_T)} \left(\sum_{\gamma \in \Gamma_T} \mu_\gamma \right) = \mu_0,$$

where Γ_T is the set of closed geodesics of length smaller than T .

Moreover, the following theorem counts asymptotically number of closed geodesics.

Theorem 5.3.2 [MARGULIS] *Let Γ_T be the set of closed geodesics of length smaller than T . Then*

$$\lim_{T \rightarrow \infty} 2T e^{-T} \#(\Gamma_T) = 1.$$

5.4 Harmonic functions

We finally move to the last topic of these notes. We encourage strongly the reader to have a look at N. Bergeron beautiful set of notes on the Laplacian on hyperbolic surfaces. We say a function f on S is an *eigenfunction* of the Laplacian of eigenvalue λ if $\Delta(f) = \lambda f$. The multiplicity of the eigenvalue λ is the dimension of the space of eigenfunctions. The Laplacian is defined locally as in the hyperbolic plane. Eigenfunctions could also have been defined just using integration on small discs and balls. Since S is assumed to be compact a general theorem asserts that the sum of the multiplicity of eigenvalues less than a given value is finite.

5.4.1 The trace formula

The length of closed geodesics and the eigenvalues of the Laplacian are related by many deep results. *Selberg trace formula* is certainly the most striking. Selberg trace formula is a generalisation of Poisson summation formula, it reads

Theorem 5.4.1 [SELBERG-TRACE VERSION] *Let S be a compact hyperbolic surface and h be an even test function satisfying some restriction. Let $\{\lambda_n\}$ be the set of eigenfunctions of the Laplacian. Let $\mu_n^2 + 1/4 = \lambda_n$, with either the imaginary part or the real part of μ_n is positive. Let $\mathcal{G}$ be the set of closed geodesics and $\ell(\gamma)$ be the length of the closed geodesic γ and $m(\gamma)$ be the multiplicity of γ . Then*

$$\sum_{n=0}^{+\infty} h(\mu_n) = -\frac{\chi(S)}{2} \int_{-\infty}^{+\infty} h(s) s \tanh(\pi s) ds + \sum_{\gamma \in \mathcal{G}} R_\gamma \hat{h}(\ell(\gamma)),$$

where

$$R_\gamma = \frac{\ell(\gamma)}{m(\gamma) 2 \sinh(\ell(\gamma)/2)}, \quad \hat{h}(s) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} h(u) e^{-ius} du.$$

We now state it using a slightly non standard approach due to Cartier and Voros. We need to introduce two generalised zeta functions.

1. The *generalised ζ function* reflects the analytic side. It is defined as

$$\zeta(s, a) = \text{Tr}(\Delta_S + a)^{-s} := \sum_{i=0}^{\infty} \frac{1}{(\lambda_n + a)^s},$$

where (λ_n) is the set of eigenvalues repeated with multiplicities.

2. The *Selberg zeta function* reflects the dynamical side. Let $\mathcal{P}$ be the set of *primitive* closed geodesics. Let us define

$$z_S(s) = \prod_{\gamma \in \mathcal{P}} \prod_{k=0}^{\infty} \left(1 - e^{\ell(\gamma)(k+s)}\right).$$

where λ_n is the set of eigenvalues of the Laplacian on S .

Then we have after taking the analytic continuation of these functions.

Theorem 5.4.2 [SELBERG-DETERMINANT VERSION] *Let S be a compact hyperbolic surface. Then*

$$\frac{\partial}{\partial s} \Big|_{s=0} \zeta_S \left(s, u^2 - \frac{1}{4} \right) = \psi(u)^{\chi(S)} z_S \left(\frac{1}{2} + u \right),$$

where $\psi(u)$ is an explicit function only depending on u , which can be interpreted as related to a spectral problem on the two-sphere. The left hand side term is usually interpreted as a regularised determinant.

5.5 Conjectures

We can now state two very famous conjectures on hyperbolic surfaces, and a slightly less one which is still interesting. The *quantum unique ergodicity conjecture* is the quantum pendant of the equidistribution of the orbits of the geodesic flows. This conjecture due to Sarnak claims the following

Conjecture 5.5.1 [SARNAK] *Let $\{\phi_n\}$ be a sequence of eigenfunctions of the Laplacian on a compact hyperbolic surface S , such that the corresponding eigenvalues go to infinity. Let μ be the hyperbolic measure on S , then for all continuous function f , we have*

$$\lim_{n \rightarrow \infty} \frac{1}{\int_S |\phi_n|^2 d\mu} \int_S f |\phi_n|^2 d\mu = \int f d\mu.$$

This conjecture is known to be "almost surely true" by Schnirelman, Zelditch and Colin de Verdière as the *quantum ergodicity theorem* in the sense that it converges for a – and then plenty – subsequence of density 1 in the sequence of functions. The following breakthrough has been obtained recently

Theorem 5.5.2 [E. LINDENSTRAUSS] *Let S be an arithmetic surface. Then the quantum ergodicity conjecture holds if we furthermore assume the sequence of functions are Hecke eigenfunctions.*

The approach is ergodic and uses the extra dynamics coming from Hecke correspondences on a space related to the solenoid.

Conjecture 5.5.3 [SELBERG] *The first eigenvalue of the Laplacian on $\mathbb{H}^2/\Gamma_0(N)$ is greater than $1/4$.*

Conjecture 5.5.4 [EHRENPREIS] *For any ϵ , for any pair of compact hyperbolic surfaces, there exists a common covering such that the two induced distances are $1 + \epsilon$ bi-Lipschitz equivalent.*

This conjecture has a nice dynamical translation on the Teichmüller space of the solenoid.