Reading Report for McShane's Identity

Junming Zhang

School of Mathematical Sciences, Nankai University
Email: junmingzhang@mail.nankai.edu.cn

June 16, 2022

Contents

1 Introduction **1**

2 Preliminaries **2**
2.1 Basic Definitions and Results 2
2.2 Collar Lemma for Cusp 3
2.3 Some Results Related to Cut Surface 6

3 Ends of Simple Geodesics **10**
3.1 A Metric for the Set of Ends of Simple Geodesics 10
3.2 The Gap Corresponding to a Closed Simple Geodesic 10
3.3 Approximating Generic Ends 11
3.3.1 Spirals into a closed geodesic 11
3.3.2 Spirals into a minimal lamination which is not a closed geodesic 13

4 McShane's Identity **18**
4.1 Structure of G_{cusp} 18
4.2 Results 19

1 Introduction

In this reading report, we give an outline of the proof of McShane's identity (cf.[3]), which is

Theorem 1.1. *On a once-punctured torus $\mathcal{S}_{1,1}$ with complete hyperbolic structure,*

$$\sum_{\gamma} \frac{1}{1+\exp|\gamma|} = \frac{1}{2},$$

where the sum takes over all closed simple geodesics γ on $\mathcal{S}_{1,1}$ and $|\gamma|$ denotes the length of γ.

Different from the higher dimension case, there are lots of complete hyperbolic structures on $\mathcal{S}_{1,1}$, so the above identity is indeed a series constant over the moduli space of these hyperbolic structures. As a generalization, G. McShane proves the following identity in [4].

Theorem 1.2. *In a finite area hyperbolic surface $\mathcal{S}$ with cusps and without boundary,*

$$\sum_{(\alpha,\beta)} \frac{1}{1+\exp\frac{1}{2}(|\alpha|+|\beta|)} = \frac{1}{2},$$

where the sum is over all unordered pairs of simple closed geodesics (α, β) (where α or β might be a cusp treated as a simple closed geodesic of length 0) on $\mathcal{S}$ such that α, β bound with a distinguished cusp point an embedded pair of pants on $\mathcal{S}$.

One recovers Theorem 1.1 from it by noting that for such embedded pants the two boundary geodesics are equal. However, we will not prove Theorem 1.2 in this reading report.

2 Preliminaries

In this reading report, unless otherwise specified, we only consider the surface equipped with a complete hyperbolic structure of finite area, and when we say punctured surface we mean a surface with at least one cusp.

2.1 Basic Definitions and Results

Definition 2.1 (lamination). *A* ***lamination*** *on a surface $\mathcal{S}$ is a closed subset of $\mathcal{S}$ which is the union of a collection of disjoint simple geodesics on $\mathcal{S}$. A non-empty lamination is* ***minimal*** *if no proper subset is a lamination. Any simple geodesic in the collection is called a* ***leaf*** *of the lamination. We say a geodesic γ* ***spirals*** *to a lamination Ω if Ω is contained in the closure of γ.*

Example 2.1. *Any closed simple geodesic is a minimal lamination.*

Example 2.2 (flat torus). *On the flat torus $\mathbb{R}^2/\mathbb{Z}^2$, the whole surface is a minimal lamination, which can be viewed as the union of all of the straight line with slope an irrational number.*

Although in the example above we get a full-measured minimal lamination, it is well known that the pointset of a geodesic lamination on a hyperbolic surface has measure zero.

By considering the ends of simple geodesic, we can classify the simple geodesics.

Lemma 2.1 (classes of simple geodesic). *On a surface any complete simple geodesic falls into exactly one of the following three classes:*

(1) It is a leaf of a compact lamination.

(2) It has a single end spiralling into a compact lamination and the other end up a cusp.

(3) It has both ends up a cusp.

Its proof can be found in [2].

To characterize how to "approximate" a minimal lamination by closed simple geodesics, we introduce the Chabauty topology on $C(\mathcal{S})$, which denotes the set consists of all closed subsets of $\mathcal{S}$.

Definition 2.2 (Chabauty topology). *Given a topological space X, the* ***Chabauty topology*** *on $C(X)$ (the set of all closed subsets of X) has a sub-basis given by sets of the following form:*

(1) $O_1(K) = \{A \mid A \cap K = \emptyset\}$, where K is compact;

(2) $O_2(U) = \{A \mid A \cap U \neq \emptyset\}$, where U is open.

If X is compact and metrizable, the Chabauty topology indeed agrees with the topology induced by the Hausdorff metric. Although the punctured surface is not compact, the convergence under this topology is still easy to describe.

Lemma 2.2 (geometric convergence). *Suppose X is a locally compact metric space. A sequence $\{A_n\}$ of closed subsets of X converges in $C(X)$ to the closed subset A if and only if:*

(1) If $\{x_{n_k}\} \in \{A_{n_k}\}$ *converges to* $x \in X$ *then* $x \in A$*;*

(2) If $x \in A$*, then there exists a sequence* $\{x_n\}$*, where each* x_n *is an element of* A_n*, converging to* x*.*

The lemma below (also proved in [2]) shows that any minimal lamination can be approximated by simple geodesics.

Lemma 2.3 (approximating minimal laminations)**.** *Let* Ω *be a minimal lamination on a complete hyperbolic surface. Then* Ω *can be approximated in the Chabauty topology by closed simple geodesics.*

To characterize the convergence of complete surfaces with basepoint, we need to introduce the geometric topology.

Definition 2.3 (ε-relation)**.** *Let* (X, x_0) *and* (Y, y_0) *be two compact metric spaces with basepoint. An* ε*-relation between* (X, x_0) *and* (Y, y_0) *is a relation* R *with the following properties:*

(1) $x_0 R y_0$*;*

(2) For each $x \in X$*, there exists* $y \in Y$ *such that* xRy*;*

(3) For each $y \in Y$*, there exists* $x \in X$ *such that* xRy*;*

(4) If xRy *and* $x'Ry'$*, then* $|d_X(x, x') - d_Y(y, y')| \leqslant \varepsilon$*.*

Definition 2.4 ((ε, r)-related)**.** *Two metric spaces with basepoint* (X, x_0) *and* $(Y, y_0)b$ *are* (ε, r)*-related if there is an* ε*-relation between compact subspaces* (X_1, x_0) *and* (Y_1, y_0) *of* (X, x_0) *and* (Y, y_0) *respectively, where* $B_X(x_0, r) \subset X_1$*, and* $B_Y(y_0, r) \subset Y_1$*.*

Definition 2.5 (geometric topology)**.** *The geometric topology on the space of isometry classes of complete locally compact path spaces with basepoint is generated by the subbasis* $\mathcal{N}(X, x_0, r, \varepsilon)$*, where* (X, x_0) *is a complete locally compact path space with basepoint ,*$r, \varepsilon > 0$*, consists of complete locally compact path spaces with basepoint* (Y, y_0) *such that* (Y, y_0) *is* (ε, r)*-related to* (X, x_0)*.*

Remark 2.1. *There is also a much more geometric viewpoint to understand the geometric topology, which is induced from the quotient of the Chabuaty topology of* $\mathsf{Isom}^+(\mathbb{H}^2)$*, see [2].*

2.2 Collar Lemma for Cusp

The famous "collar lemma" tells us every closed simple geodesic have an embedding tubular neighborhood. Similarly for every cusp on the surface, it also has a "collar".

Definition 2.6 (cusp region)**.** *A portion of the surface* $\mathcal{S}$ *isometric to*

$$\{z \in \mathbb{H}^2 \mid \operatorname{Im} z \geqslant 1\}/[z \mapsto z + p]$$

*is called a **cusp region** of length (of the bounding horocyclic curve)* p*.*

Theorem 2.1 (collar lemma for cusp)**.** *On a surface $\mathcal{S} := \mathcal{S}_{g,n}$ (genus equal to g with n cusps), where $g, n \geqslant 1$, every cusp has a neighborhood which is a cusp region of length 2.*

Proof. We can cut along two closed geodesics (one closed geodesic on $\mathcal{S}_{1,1}$) on $\mathcal{S}$ to get an embedded pants contained the cusp. Now we can cut the pants along the geodesic from the cusp to two boundary components and the geodesic connecting two boundary components to get two congruent pentagons with an ideal vertex and four right angles. Then we can put one of it in $\mathbb{H}^2$ with two infinite sides are portions $\{z \mid \operatorname{Re} z = 0\}$ and $\{z \mid \operatorname{Re} z = 1\}$ respectively. Then another three sides are three arcs between them. One can easily observe thar the radii of these three arcs are all smaller than 1, so this pentagon contains a horocyclic curve, $\{z \mid \operatorname{Re} z \in [0,1], \operatorname{Im} z = 1\}$, centered at the cusp, ∞. So $\{z \mid \operatorname{Re} z \in [0,1], \operatorname{Im} z \geqslant 1\}$ is embedded in to $\mathcal{S}$ Then we can reglue these two pentagons to get the cusp neighborhood we need.

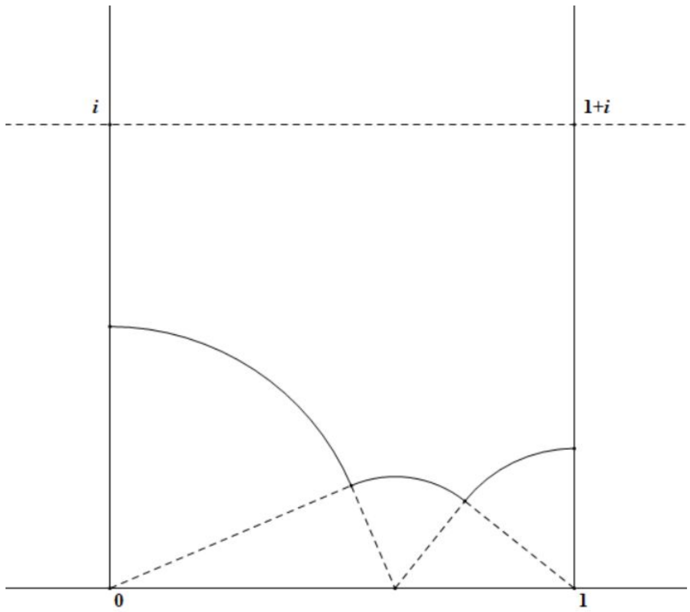

Figure 1: collar lemma for cusp

□

In fact, every simple geodesic goes into the cusp region must have very good behavior.

Lemma 2.4 (behavior of simple geodesic intersects the cusp region)**.** *Let $\mathcal{S}$ be a surface with a cusp region of length 2. If a simple complete geodesic intersects with the cusp region whose bounding curve has length less than 2, it must have an end up to the cusp and intersect the bounding curve perpendicularly.*

Proof. Lift the surface to $\mathbb{H}^2$, such that the cusp region whose bounding curve has length 2 is covered by $\{z \mid \operatorname{Im} z \geqslant 1\}/[z \mapsto z+2]$. Suppose the simple geodesic is not perpendicular to $\{z \mid \operatorname{Im} z = 1\}$, then we can lift the simple geodesic to a semicircle with radius larger than 1, now it must intersect its another lifting which induced by $[z \mapsto z+2]$. So it is not simple, contradiction.

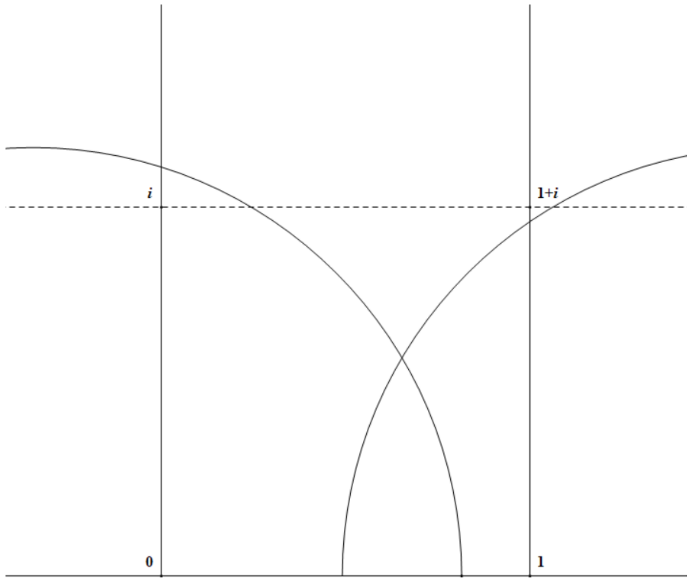

Figure 2: behavior of simple geodesic intersects the cusp region

□

As a corollary, we can deduce that

Corollary 2.1. *Let $\varepsilon > 0$. Any geodesic which is the leaf of a compact lamination on a punctured surface does not intersect the cusp region whose bounding curve has length $2 - \varepsilon$.*

Proof. By Lemma 2.4 a complete simple geodesic which intersects the cusp region whose bounding curve has length $2 - \varepsilon$ must have an end up the cusp. Thus it cannot be contained in any compact subset of the surface. □

and this implies that

Corollary 2.2. *There exists point on the punctured surface $\mathcal{S}$ such that it is disjoint from all simple closed geodesics.*

Proof. Any point in the cusp region of length smaller than 1 is an required point. □

Remark 2.2. *This corollary has a much stronger version, which is proved by Birman and Series in [1], that is, let G be the set of all simple geodesics on a hyperbolic surface, the set S of points which lie on a geodesic $\gamma \in G$ has Hausdorff dimension* 1*, hence zero measure.*

and

Corollary 2.3 (simple geodesics disjoint in cusp region)**.** *Let γ be a complete simple geodesic on a punctured surface. The geodesic γ intersects no other complete simple geodesic in any cusp region whose bounding curve has length less than* 2*.*

For the once-punctured torus $\mathcal{S}_{1,1}$, we can do much better.

Theorem 2.2 (collar lemma for cusp on $\mathcal{S}_{1,1}$)**.** *The cusp on the once-punctured torus $\mathcal{S}_{1,1}$ has a neighborhood which is a cusp region of length 4. Let $\varepsilon > 0$. Any minimal compact lamination on $\mathcal{S}_{1,1}$ does not intersect the cusp region whose bounding curve has length $4 - \varepsilon$.*

Proof. The proof is similar to Theorem 2.1, we still cut the surface along a closed geodesic to get a pair of pants and then cut along some geodesics connecting the cusp and boundary components to get two congruent pentagons. Now since the two boundary components are equal, we can cut these pentagons into four congruent quadrilaterals with one ideal vertex and three right angles. Then we can put one of them into $\mathbb{H}^2$ such that its two infinite sides are portions $\{z \mid \mathrm{Re}\, z = 0\}$ and $\{z \mid \mathrm{Re}\, z = 1\}$ respectively. Now one can observe that the Euclidean radius of the two finite sides are still smaller than 1. So we can reglue these four quadrilaterals to get a cusp region of length 4. And from the above discussion we see that the chosen closed simple geodesic is away from this region. Since every compact minimal lamination is the limit of some closed geodesics under the Chabauty topology, we get the required conclusion.

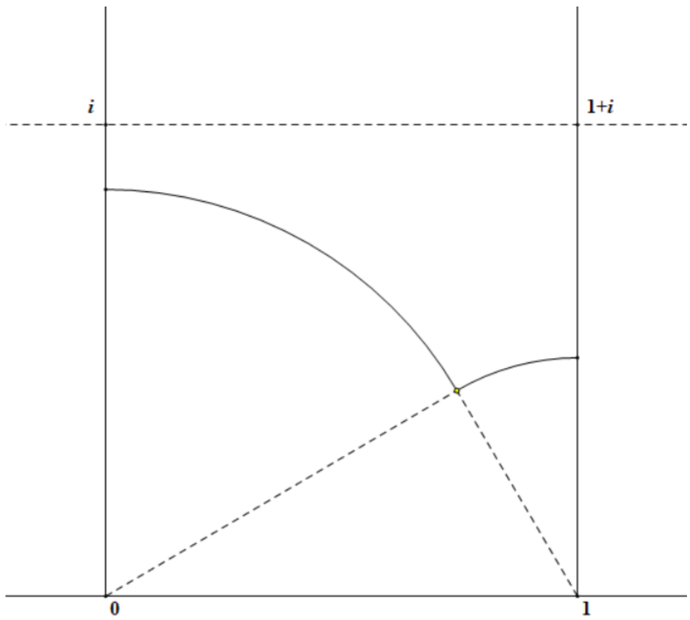

Figure 3: collar lemma for cusp on once-punctured torus

□

2.3 Some Results Related to Cut Surface

Definition 2.7 (cut surface)**.** *Choose a disjoint set of simple curves for which the underlying pointset is closed as a subset of the surface. The complement of this set of curves in the surface is an open surface. The* ***cut surface*** *is the metric completion of this open surface with respect to its path metric.*

If we choose these simple curves to be geodesics, the cut surface we get have geodesic boundaries.

Example 2.3. *When we cut $\mathcal{S}_{1,1}$ along a closed geodesic the surface we get is a pair of pants.*

There are compact minimal laminations on the punctured torus which are not closed geodesics, although they are the limits of closed simple geodesics (in the Chabauty topology on closed subsets). The surface we obtain when we cut along such a lamination is not a pair of pants, but a limit of pairs of pants. For instance, it is a punctured ideal bigon.

Lemma 2.5 (cut lamination not a closed geodesic)**.** *If $\mathcal{S}_{1,1}$ is cut along a minimal compact lamination which is not a closed geodesic then the cut surface is a punctured ideal bigon. That is, it is isometric to an ideal triangle doubled along two of its edges.*

Proof. Let Ω be a minimal lamination which is not a closed geodesic. By Lemma 2.3, we can choose a sequence of closed geodesics γ_i, which converge to Ω in the Chabauty topology, so that the lengths are monotone increasing. We cut along γ_i to obtain a sequence of pairs of pants.

We choose a point on the torus disjoint from all simple closed geodesics (cf. Corollary 2.2). The image of this point on each of the pants is a base point. The sequence of pants with this base point converges to the double of an ideal triangle in the geometric topology (the lengths of the boundary components increases monotically and the length of the simple geodesic arcs joining these boundaries decreases monotically.)

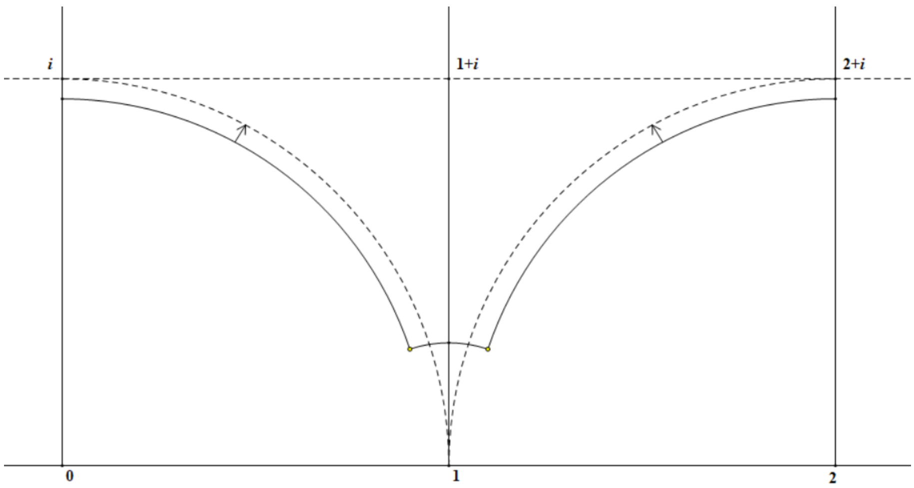

Figure 4: cut lamination not a closed geodesic

□

Lemma 2.6 (minimal laminations intersect)**.** *Two distinct minimal compact laminations on $\mathcal{S}_{1,1}$ intersect.*

Proof. Let Ω be a minimal compact lamination. We cut along Ω and analyse the cut surface we obtain. If the torus has a lamination disjoint from Ω then there must be a corresponding lamination on this cut surface.

Suppose Ω is a closed geodesic. Cutting along Ω we obtain a pair of pants. The only closed simple geodesics on a pair of pants are contained in the boundary (recall that every free homotopy class on a hyperbolic surface has a unique geodesic represents it). So there are no simple closed geodesics on the torus which are disjoint from Ω. In addition the simple closed geodesics are dense in minimal laminations (cf. Lemma 2.3), so the cut surface has no minimal laminations other than those contained in the boundary. Thus there are no minimal laminations disjoint from Ω.

Suppose Ω is not a closed geodesic. When we cut along it now the cut surface is a punctured ideal bigon. This surface has no closed simple geodesics and so can have no minimal laminations. □

Corollary 2.4 (disjoint from unique minimal)**.** *On a once-punctured torus a simple geodesic with a single end up the cusp is disjoint from exactly one minimal lamination. The geodesic spirals into this lamination.*

Proof. Let γ be a simple geodesic with a single end up the cusp. If γ intersects the minimal lamination in its closure then it must self intersect. So it is disjoint from at least one lamination. However, by Lemma 2.6 any other compact lamination intersects this one. So γ must intersect every other compact lamination on the punctured torus. □

Lemma 2.7 (number of disjoint geodesics)**.** *On a once-punctured torus:*

(1) A closed simple geodesic is disjoint from exactly one simple geodesic with both ends up the cusp and vice versa. The closed simple geodesic is also disjoint from exactly four simple geodesics which each have a single end up the cusp.

(2) A compact minimal lamination which is not a closed geodesic meets every geodesic with both ends up the cusp. It is disjoint from exactly two simple geodesics which each have a single end up the cusp.

Proof. Suppose the lamination is a simple closed geodesic. The cut surface is a pair of pants. There is a unique free homotopy class of simple curves on the surface with both ends up the cusp and so there must be a unique simple geodesic with both ends up the cusp. Every closed simple curve on this surface is homotopic either to the cusp or to one of the boundary components. So there are no compact minimal laminations on the surface other than those contained in the boundary. Fix a boundary component. There are exactly two simple geodesics which each have an end up the cusp and another end spiralling to this boundary component.

For instance, suppose α is the simple closed geodesic and γ is a simple geodesic with two ends up to the cusp intersects with it (e.g. the geodesic generated from gluing two geodesics from the cusp to two boundary components on the cut surface along α). We lift to $\mathbb{H}^2$ such that γ' is a lift of γ which is a vertical line. Travelling from the cusp at infinity, let F be the first leaf γ' meets of the lamination consists of the lifting of α, and let α_1 and α_2 be the pair of vertical lines which end at the endpoints of F. The image of α_1 on the surface is a simple geodesic, since we can find a simple curve on the surface which with a lift to $\mathbb{H}^2$ which has the same endpoints as the lift of α_1, namely the curve from the cusp along γ' to $\gamma' \cap F$, then goes along F to the endpoint of α_1 (this curve projects to a simple curve since F is the first leaf γ' meets so no other lifting of this curve intersects the curve on $\mathbb{H}^2$). Similarly the image of α_2 on the surface is simple. And they both spiral into α since they both have common end points with F.

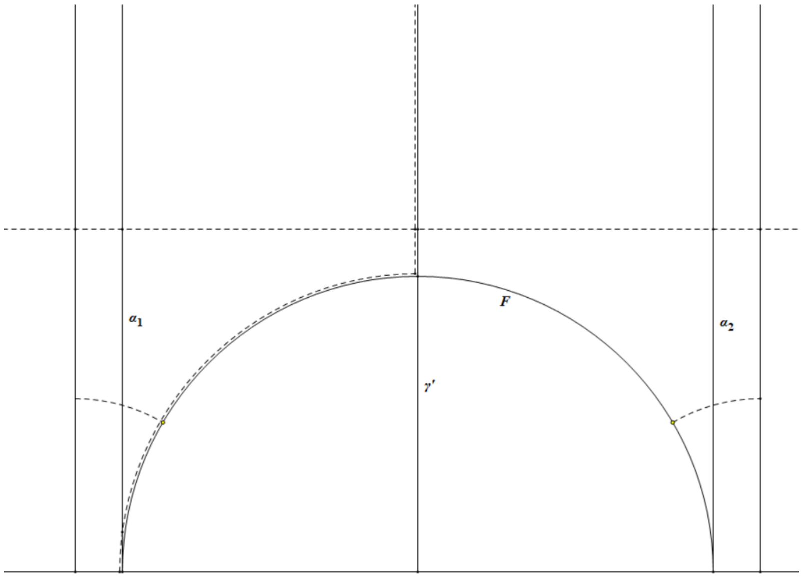

Figure 5: number of disjoint geodesics

So there are four geodesics with a single end up the cusp disjoint from the geodesic we cut along. For any simple geodesic with both ends up the cusp, by cutting along it and considering the free homotopy class we can get the converse proposition.

Suppose the lamination is compact and minimal but not a closed geodesic. The cut surface is a punctured ideal bigon. There is no free homotopy class of simple closed curve with both ends up the cusp on this surface. We think of the bigon as being the double of an ideal triangle along two of its edges. In this way we see that there are two simple geodesics each with a single end up the cusp, namely the geodesic edges of the triangle along which we doubled. □

3 Ends of Simple Geodesics

3.1 A Metric for the Set of Ends of Simple Geodesics

Let $\mathcal{S}_{1,1}$ be the punctured torus and C be the cusp region whose horocyclic boundary curve ∂C has length smaller than 2. The intersection of C with a complete simple geodesic is empty, or has one component, or has two components. We denote by G_{cusp} the set of all such components. This set can be identified with a certain closed subset of C since the set of geodesic laminations is closed in the space of all closed sets under the Chabauty topology. Note that a complete simple geodesic with two ends up the cusp will give rise to two points of G_{cusp}.

Definition 3.1. *Let α and β be geodesics in G_{cusp}. We define the **distance** between α and β, $d_{\text{cusp}}(\alpha, \beta)$, to be the minimum distance measured by arclength along ∂C between the sets $\alpha \cap \partial C$ and $\beta \cap \partial C$, divided by the total length of ∂C.*

Remark 3.1. *This definition is independent of the choice of cusp region due to Lemma 2.4.*

One can also define this metric by using the area of the strip between two ends in G_{cusp} from the easy calculation below.

Definition 3.2 (strip)**.** *A **strip** is a component of the complement in the cusp region C of some (non-empty) closed collection of geodesics in G_{cusp}.*

Lemma 3.1 (area of strip)**.** *The area of a strip is just the length of its horocyclic side.*

Proof. Note that any strip is congruent to the region $\{z \mid \operatorname{Im} z \geqslant 1, 0 < \operatorname{Re} z < p\}$ for some $p > 0$. Then the area of the strip is

$$\int_1^\infty \int_0^p \frac{\mathrm{d}x\mathrm{d}y}{y^2} = p,$$

which equals to the length of its horocyclic side. □

3.2 The Gap Corresponding to a Closed Simple Geodesic

From Lemma 2.7, we know that every closed simple geodesic on $\mathcal{S}_{1,1}$ corresponds to a unique simple geodesic with two ends up to the cusp and vice versa. In this section, we will show that the two points corresponding to this geodesic are isolated in G_{cusp} and compute how far do they leave from other points.

Definition 3.3. *Let G_{one} be the set of ends of all complete simple geodesics each with a single end up the cusp and let G_{two} be the set of all ends of complete simple geodesics with both ends up the cusp. A **gap** is a maximal strip in the complement of G_{one} in the cusp region.*

Lemma 3.2 (spiralling-in closer)**.** *Suppose $\gamma, \beta \in G_{\text{cusp}}$. Let α be a closed simple geodesic which intersects γ and is such that β neither intersects a nor spirals into α. There is a geodesic α' in G_{cusp}, with a single end up the cusp and the other end spiralling to α, such that*

$$d_{\text{cusp}}(\alpha', \beta) < d_{\text{cusp}}(\gamma, \beta).$$

Proof. We define γ', F, α_1, α_2 same as Lemma 2.7, and suppose the horocycle of length 1 lifts to $\{z \mid \operatorname{Im} z = 1\}$ in addition. We show that one of α_1, α_2 is nearer to β than γ is.

The horocyclic segment L contained in $\{z \mid \operatorname{Im} z = 1\}$ which abuts on α_1 and α_2 maps injectively to a horocyclic segment for the following reasons. The line $\{z \mid \operatorname{Im} z = 1/2\}$ is a lift of the horocycle of length 2 on the surface. Since α is simple and closed it cannot intersect the horocycle of length 2, by Lemma 2.4. So the lift F of α cannot intersect $\{z \mid \operatorname{Im} z = 1/2\}$. Therefore the Euclidean diameter of the semicircle F is less than 1. So L projects to something of length less than 1 on the horocycle of length 1 on the surface.

We will be done if no lift of β can intersect L. To see this note that α_1 is not a lift of β because the image of β on the surface cannot spiral into α. For the same reasons α_2 is not a lift of β. If there were a lift of β which met L then it would intersect L orthogonally by Lemma 2.4. Hence it would meet F, which is impossible. ☐

Theorem 3.1 (existence of gap and its width)**.** *Let β be a simple geodesic with both ends up the cusp. There are a pair of strips such that β is the only simple geodesic which intersects either of these strips. The width of each such strip is*

$$\frac{1}{1 + \exp|\alpha|}$$

where α is the unique closed simple geodesic disjoint from β.

Proof. The existence of such strip follows from Lemma 2.7 and Lemma 3.2 directly. To compute its width, we only need to compute the distance in G_{cusp} between β and the geodesic α_1 (or α_2) we construct in Lemma 3.2, and the width of such a strip is twice the value, and one can complete this calculation readily. ☐

3.3 Approximating Generic Ends

We would like to prove that G_{one} has no isolated point in this section. There are two classes of points in G_{one}:

(1) the other end spirals into a closed geodesic;

(2) the other end spirals into a minimal lamination which is not a closed geodesic.

We will approximate these two classes by different methods.

3.3.1 Spirals into a closed geodesic

Lemma 3.3 (curve that cuts only once)**.** *Let γ be a simple geodesic on $\mathcal{S}_{1,1}$ with one end up the cusp and the other end spiralling to a closed geodesic Ω. There is a complete simple geodesic, β, such that*

(1) It meets Ω in a single point.

(2) For any lift γ' of γ, there is a lift Ω' of Ω and a lift, β' of β such that γ', β', Ω' form a triangle in $\mathbb{H}^2$.

Proof. Consider the two geodesics from the cusp to the boundary components on the cut surface along Ω, then we can reglue them to get a complete geodesic satisfying the condition. □

Remark 3.2. *This lemma also holds for any punctured surface without boundary, but we need to consider whether Ω separates the surface, see [3] or [4].*

We now define a Dehn twist.

Definition 3.4 (Dehn twist)**.** *Let γ be a closed simple geodesic on a surface. By Magulis's lemma, there is a neighborhood of γ which is homeomorphic to an annulus. Parameterize this neighborhood as*

$$\{(r,z) \mid 1 \leqslant r \leqslant 2, z \in S^1\}.$$

*The **Dehn twist** round γ is the homeomorphism which is the identity out this annulus and the map $(r,z) \mapsto (r, \mathrm{e}^{2\pi \mathrm{i} r} z)$ on the annulus.*

The image of a geodesic β which intersects γ, under a Dehn twist round γ, is not a geodesic. However, it is a piecewise smooth curve and there is a unique geodesic determined by straightening. If the geodesic β is simple then the geodesic determined by its image is simple. If the geodesic β intersects γ exactly n times then the geodesic determined by the image of β under a Dehn twist round γ intersects γ exactly n times.

Theorem 3.2 (approximate by Dehn twist)**.** *Let γ be a simple geodesic with one end up the cusp and the other end spiralling to a closed geodesic Ω. Then there is a sequence of points of G_{cusp} converging to γ. These points can be taken to be ends of a sequence of geodesics which are determined by the images of a certain geodesic under iterated Dehn twists round Ω. The choice of this geodesic depends on γ.*

Proof. We lift to $\mathbb{H}^2$. Conjugate the group of covering transformations for $\mathcal{S}_{1,1}$ so that the line $\gamma' = \{z \mid \operatorname{Re} z = 0\}$ is a lift of γ, and the semi-circle, Ω', with endpoints 0 and 1 is a lift of Ω. Let A be the hyperbolic transformation which covers this lift. We may suppose that the attracting fixed point of A is 0. Let β be the geodesic constructed in Lemma 3.3 and β' is the corresponding lift. Since β intersects γ only once on the surface β' intersects exactly one leaf of the lamination of $\mathbb{H}^2$ by lifts of γ. Let T denote the Dehn twist along Ω.

Let $\beta'_0 = \beta'$ and let β'_n, for $n \geqslant 1$, be the vertical line with endpoint the image of the finite endpoint of β' under A^n. One shows that β'_n covers $T^n\beta$ as follows. In the surface, one can construct a piecewise geodesic curve in the same homotopy class as $T^n\beta$ by following β down from the cusp till it meets Ω, then going n times round this closed geodesic (in the right direction) and finally returning to the cusp via β. In th universal cover $\mathbb{H}^2$ this construction corresponds to following β' down from ∞, walking along Ω' until we come to the n-th lift of β that we meet, and then travelling along this lift of β to its endpoint on the other side of Ω' from ∞. Since the n-th lift of β that we meet will be precisely $A^n\beta'$ and the endpoint concerned is the image of the finite endpoint of β' under A^n. Thus β'_n has the same endpoints as a lift of a curve in the same homotopy class as $T^n\beta$ and since β'_n is a geodesic, it must cover $T^n\beta$.

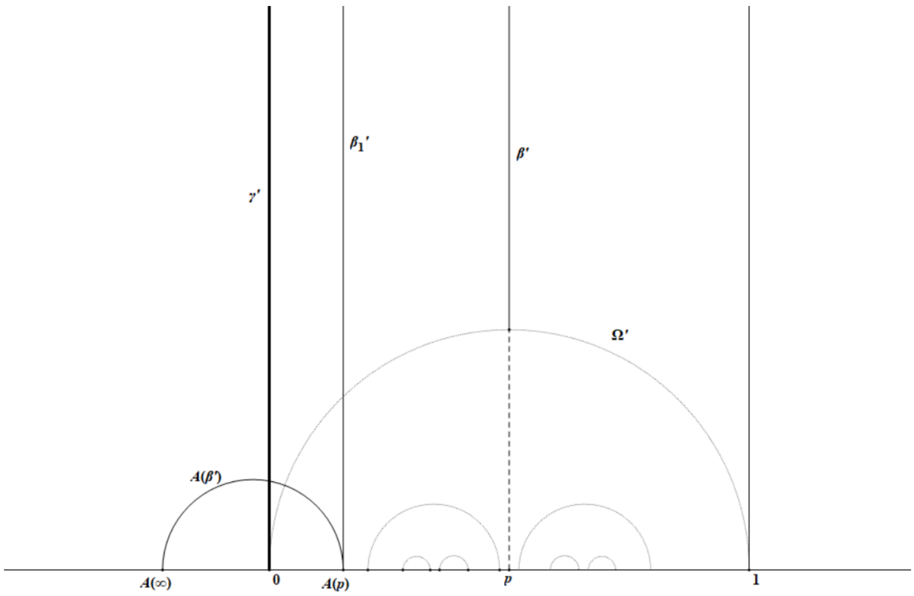

Figure 6: approximate by Dehn twist

The sequence of finite endpoints of the β_n' converges to the attracting fixed point of A which is none other than the finite endpoint of γ'. Thus β_n converges to γ' and so, when we take the projection to the quotient surface, $T^n\beta$ converges to a lamination containing γ. $\square$

3.3.2 Spirals into a minimal lamination which is not a closed geodesic

We first fix some notations. When we work in the universal cover of a surface we will identify the cover with the upper half plane. Throughout this section γ will be a simple geodesic with a single end up the cusp. By $\gamma_{\min}$, we mean the minimal lamination that γ spirals into. It will be a running hypothesis that $\gamma_{\min}$ is not a closed geodesic.

For our convenience, we give the surface an orientation. The geodesic γ has a single end up the cusp. For a point on γ we introduce a notion of left and right by insisting that the observer face along γ looking towards this end.

With this notion we define the leaf of $\gamma_{\min}$ nearest γ on the right as follows. Since $\gamma_{\min}$ is not a closed geodesic, the corresponding cut surface has no boundary component which is a circle. The boundary of the cut surface consists of a number of components which are all doubly infinite geodesics. There are exactly two such components in the case of a punctured torus. These components bound portions of the surface called **crowns**. Each crown consist of a number of **spikes**, which are portions of the surface isometric to $\{z \mid 0 \leqslant \operatorname{Re} z \leqslant 1, \operatorname{Im} z \geqslant p\}$, for some $p > 0$, with the geodesic sides contained in the boundary of the cut surface. For an observer on y, who is on the image of this spike on the closed surface, the leaf of $\gamma_{\min}$, nearest γ on the right is the boundary leaf of $\gamma_{\min}$ on his right as he looks along γ towards the cusp and away from the spike.

Let β be the leaf of $\gamma_{\min}$ nearest γ on the right. Note that β is necessarily a boundary leaf of $\gamma_{\min}$. We can lift γ to a vertical line γ' in $\mathbb{H}^2$ and β to a semicircle

β' tangent to it at the finite end point of γ' in the right. We can construct a curve in the universal cover from ∞ along γ and then goes to β by a small arc, then goes to the highest point of β' along β'. If the small arc is short enough, the curve we get above projects a simple curve on $\mathcal{S}_{1,1}$ since γ is disjoint from $\gamma_{\min}$, then it gives a simple geodesic on $\mathcal{S}_{1,1}$. We denote it by $\beta^{\perp}$, which is the unique geodesic which intersects β perpendicularly in a single point, and which, together with γ and some portion of β, bounds an embedded triangle.

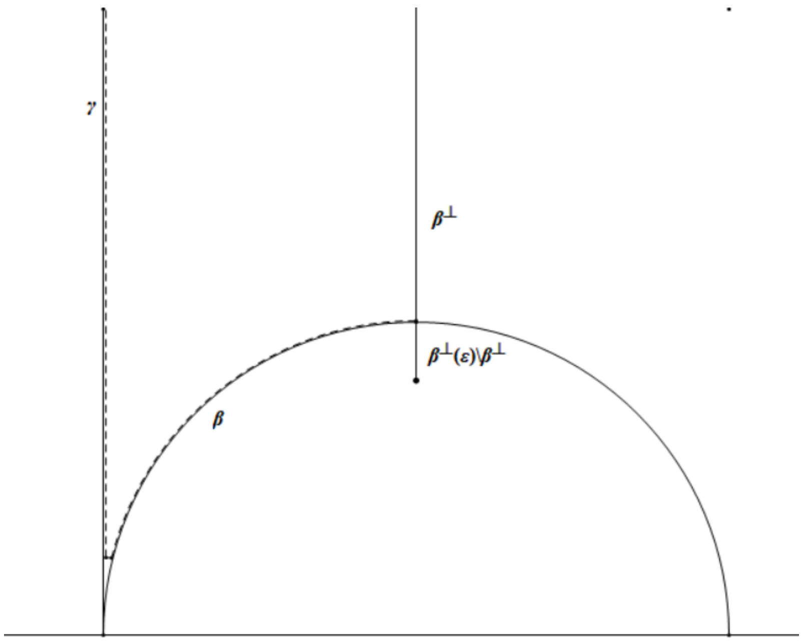

Figure 7: construction of $\beta^{\perp}$

Since the injective radius at $\beta \cap \beta^{\perp}$ is larger than 0, we can extend $\beta^{\perp}$ into a larger simple geodesic $\beta^{\perp}(\varepsilon)$ such that $\beta^{\perp}(\varepsilon) \setminus \beta^{\perp}$ has length ε. And we can choose ε small enough such that the angle made when γ intersects $\beta^{\perp}(\varepsilon)$ is sufficiently close to a right angle.

Lemma 3.4 (infinite intersections)**.** *Let $T > 0$. The intersection of the curve $\{\gamma(t) \mid t > T\}$ with $\beta^{\perp}(\varepsilon)$ is an infinite set.*

Proof. If $\{\gamma(t) \mid t > T\}$ and $\beta^{\perp}(\varepsilon)$ are disjoint then there is an open ball centre $\beta^{\perp} \cap \beta$ disjoint from $\{\gamma(t) \mid t > T\}$ since γ and β are disjoint. But this contradicts the fact that β is contained in the closure of γ. □

Definition 3.5 (height, closest approach)**.** *Let x be a point on $\beta^{\perp}(\varepsilon) \setminus \beta^{\perp}$. We define the **height** of x, $\Theta(x)$, as the distance from x to $\beta \cap \beta^{\perp}$. Since ε is small, the distance can be measured along $\beta^{\perp}(\varepsilon)$.*

*A **closest approach** for γ to the point $\beta \cap \beta^{\perp}$ is a value of the parameter T such that $\gamma(T)$ lies on $\beta^{\perp}(\varepsilon)$ and such that $\Theta(\gamma(T))$ is less than $\Theta(\gamma(t))$ for all $t < T$, such that $\gamma(t) \in \beta^{\perp}(\varepsilon)$.*

Definition 3.6 (join, double join)**.** *Let S and T be closest approaches for γ. We call the union of $\{\gamma(t) \mid t < T\}$ and the portion of $\beta^{\perp}(\varepsilon)$ which connects $\gamma(T)$ to the* *cusp the **join** and the curve obtained by following $\beta^{\perp}(\varepsilon)$ from the cusp to $\gamma(S)$ then following γ to $\gamma(T)$ then going up the cusp along $\beta^{\perp}(\varepsilon)$ the **double join** between S and T.*

We will use join and double join approximate γ below, however, we need to prove that they give simple geodesics.

Lemma 3.5. *Both join and double join give nontrivial simple geodesic.*

Proof. The only nontrivial part is that double join gives a nontrivial homotopy class on $\mathcal{S}_{1,1}$, which is equivalent to there exists a lift of the double join has two different ends, which comes from γ intersects with $\beta^{\perp}(\varepsilon)$ almost perpendicularly at closest approach.

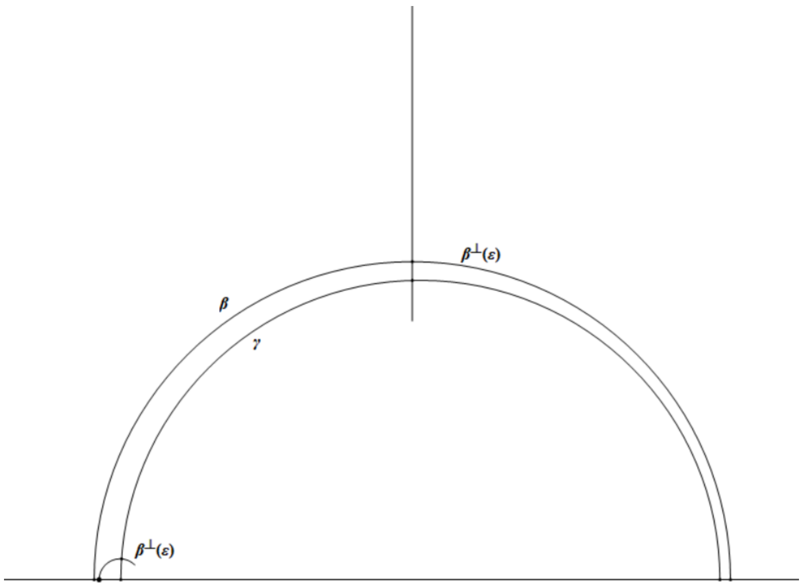

Figure 8: double join gives nontrivial simple geodesic

□

Definition 3.7 (approach from the left or right)**.** *Let α be an infinite geodesic with a distinguished end. Suppose α meets $\beta^{\perp}(\varepsilon) \setminus \beta^{\perp}$ at the point x. The intersection point x is an **approach from the left (resp. right)** if the the end of α is on the left (resp. right) side of an observer at the point x who faces along $\beta^{\perp}(\varepsilon)$ looking towards the cusp.*

Lemma 3.6 (approximation estimates)**.** *Let α be a geodesic with a distinguished end which does not intersect β or $\beta^{\perp}$ and has an endpoint x on $\beta^{\perp}(\varepsilon) \setminus \beta^{\perp}$. Suppose α is an approach from the left. The curve which is the union of α from its distinguished ideal endpoint to x, and the portion of $\beta^{\perp}(\varepsilon)$ between x and the cusp, can be straightened to a geodesic with at least one end up the cusp. There is a constant, K, depending only on β and γ, such that one end of this geodesic is closer than $K\Theta(x)$ to γ.*

Proof. We lift to $\mathbb{H}^2$ so that the horocycle of length 1 is contained in $\{z \mid \operatorname{Im} z = 1\}$ and γ lifts to the line $\gamma' = \{z \mid \operatorname{Re} z = 0\}$. There is a unique lift β' of β, which is a semicircle with endpoints 0 and p, for some $0 < p \leqslant 1/2$ (cf.Theorem 2.2). There is a lift, of $\beta^{\perp}(\varepsilon)$ contained in the line $\{z \mid \operatorname{Re} z = p/2\}$.

Under our hypothesis there is a lift α' of α lying entirely below the semicircle β' which intersects $\beta^{\perp}(\varepsilon)$.

We note that 0 as an endpoint of γ'. So the distance, as measured by the metric on G_{cusp}, between the geodesic γ and the end of the join of α and $\beta^{\perp}(\varepsilon)$ is the distance between 0 and the endpoint of α on $\mathbb{R}$.

Since α' lies entirely below β' this distance is less than the Euclidean diameter of β' minus the Euclidean diameter of α'. The Euclidean diameter of α' is greater than twice the imaginary part of any point on it, in particular greater than $2 \operatorname{Im} x'$, where x' is the intersection of $\beta^{\perp}(\varepsilon)$ and α' and it projects the intersection x of $\beta^{\perp}(\varepsilon)$ and α. Then

$$\Theta(x) = \int_{\operatorname{Im}(x')}^{p/2} \frac{\mathrm{d}y}{y} = \log(p/2) - \log \operatorname{Im}(x'),$$

so

$$\operatorname{Im} x' = \frac{p}{2} \mathrm{e}^{-\Theta(x)}.$$

Thus we can get the distance is smaller than $p(1 - \mathrm{e}^{-\Theta(x)}) \leqslant p\Theta(x)$.

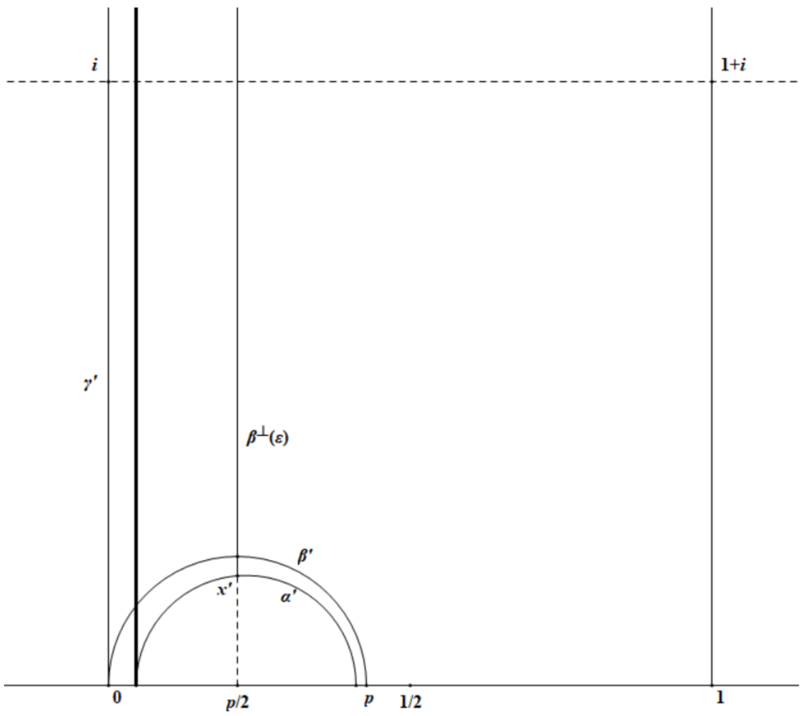

Figure 9: approximation estimates

□

We need another lemma for the following approximate.

Lemma 3.7. *Let $\varepsilon > 0$. Let γ be the geodesic $\{z \mid \operatorname{Im} z = 0\}$ and β be the geodesic with endpoints 0 and 1. Let z be the highest point on β, x the point at the same Euclidean height as z on γ, and y any point on γ.*

If x' and y' are points of $\mathbb{H}^2$ such that $d(z,x') < \varepsilon$ and $d(y,y') < \varepsilon$ then

$$d(x',y') \geqslant d(x,y) - 2\varepsilon - 1,$$

where d is the hyperbolic metric on $\mathbb{H}^2$.

Proof. The points z and x are joined by a horocyclic curve of length z so $d(z,x) < 1$. The lemma then follows from the triangle inequality. □

Theorem 3.3 (approximating from right)**.** *Let γ be a simple geodesic with a single end up the cusp and its other end spiralling into a minimal lamination which is not a closed geodesic. Then the geodesic γ can be approximated from the right by simple geodesics with both ends up the cusp.*

Proof. The geodesic γ has a single end up the cusp, we take this as its distinguished end in the manner of Definition 3.7. Let $T_1 < T_2 < T_3 < \dots$ be a complete list of the closest approaches of γ along $\beta^{\perp}(\varepsilon)$.

If there is an infinite subsequence of closest approaches, T_{i_n} which are from the left of $\beta^{\perp}(\varepsilon)$, then there is a corresponding sequence of geodesics with both ends up the cusp approximate γ from the right. Such a sequence is obtained by straighten the joins at T_{i_n} and they indeed converge to γ by Lemma 3.6.

Suppose, for a contradiction, that there exists an N such that for all $i \geqslant N$, T_i is a closest approach from the right.

Let x be the point of γ on the same horocycle centered at the cusp as $\beta \cap \beta^{\perp}$ and let $t(x)$ be the parameter value corresponding to x.

We now choose N (bigger if necessary) so that T_N is greater than $t(x) + 2\varepsilon + 1$. Let $n > N$. We lift to $\mathbb{H}^2$ such that γ is $\{z \mid \operatorname{Re} z = 0\}$. Conjugate the covering group if necessary so that there is a lift of β with endpoints $0, 1$. There is then a lift $\beta^{\perp}(\varepsilon)$ contained in $\{z \mid \operatorname{Re} z = 1/2\}$. We lift $\gamma(T_n)$ to a point on this line. This lift of $\gamma(T_n)$ lies on a unique lift of γ (different from $\{z \mid \operatorname{Re} z = 0\}$), we denote this by γ'. There is a lift of $\gamma(T_n)$, which we will also denote by $\gamma(T_{n+1})$, which is a point on γ'.

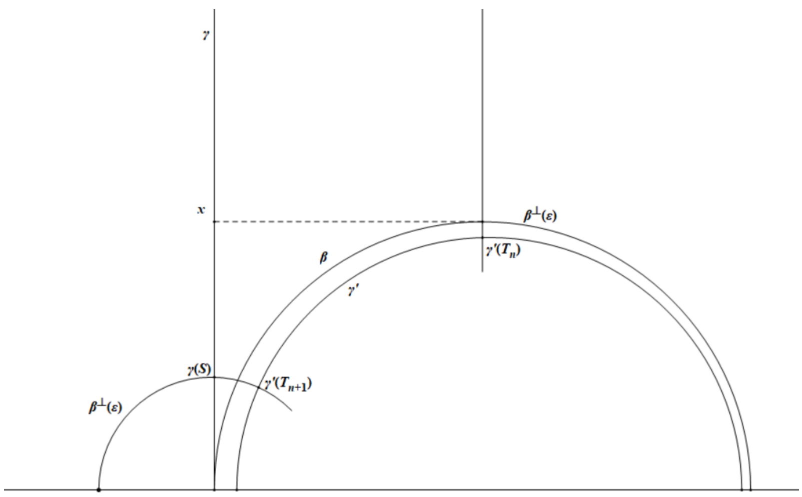

Figure 10: approximating from right

Having chosen this lift $\gamma(T_{n+1})$ there is a lift $\beta^{\perp}(\varepsilon)'$ of $\beta^{\perp}(\varepsilon)$, through this point. Since T_{n+1} is an approach from the right to $\beta^{\perp}(\varepsilon)$, this lift has its ideal endpoint to the left of both endpoints of γ'. If this endpoint were negative, then the intersection of $\beta^{\perp}(\varepsilon)'$ and γ would give a point on the surface on γ which was a closer approach than T_{n+1}. We show that this cannot happen by proving that the parameter value along γ of this intersection would be less than T_{n+1}.

Let S be the parameter value of this intersection. Since γ never intersects $\beta^{\perp}$,

$$d(\gamma(T_{n+1}), \gamma(S)) < \varepsilon.$$

By Lemma 3.7,

$$T_{n+1} - T_n > S - t(x) - 2\varepsilon - 1.$$

So, since $n > N$,

$$T_{n+1} > S + T_n - t(x) - 2\varepsilon - 1 > S + T_n - T_N > S.$$

This contradicts the hypothesis that T_{n+1} was a closest approach. So the double join between T_n and T_{n+1} approximates γ from the right. □

4 McShane's Identity

4.1 Structure of G_{cusp}

We will show every gap in G_{cusp} correspond to a closed geodesic in the sense of Theorem 3.1.

Theorem 4.1 (geodesic which bounds a gap)**.** *A geodesic arc which bounds a gap is contained in a complete simple geodesic which spirals into a closed geodesic.*

Proof. Let γ be a geodesic which bounds a gap. Since the set of geodesic laminations is closed in the Chabauty topology, the complete geodesic that contains γ must be simple (may be isolated in G_{one} or can be approximated by a sequence in G_{one}). And γ must be in G_{one} since any geodesic in G_{two}. Suppose γ spirals into a lamination which is not a closed geodesic. We apply Theorem 3.3 to get a sequence which converges to γ from the right and as was indicated in the text preceding Theorem 3.3 we can construct a sequence which converges from the left in an analogous manner. This contradicts the hypothesis that γ bounds a gap. □

Theorem 4.2 (Cantor set)**.** *G_{one} (with the induced metric from G_{cusp}) is a Cantor set, i.e. a nowhere dense closed set without isolated point.*

Proof. G_{one} is closed as a subset of G_{cusp} with its metric, since G_{two} is isolated hence open.

Let $\alpha \in G_{\text{one}}$ and Ω be the minimal lamination it spirals to. There is a sequence β_n of geodesics in G_{two} which converge to a lamination containing α (if Ω is a closed geodesic this follows from Theorem 3.2, otherwise it follows from Theorem 3.3). Then G_{one} has no isolated points follows from Lemma 3.2.

Suppose there is a point x interior to some interval contained in G_{one}. Since G_{two} is dense in G_{cusp} there is a sequence, x_n converging to x with $x_n \in G_{\text{two}}$. By Theorem 3.1 each of the x_n is interior to a gap, contradiction. □

Corollary 4.1 (geodesics which bound same gap)**.** *Let J be a gap. The two geodesics which bound J spiral into the same closed geodesic.*

Proof. Let γ be one of the geodesics which bound J. By Theorem 3.1 there is a gap K which is bounded by γ and another geodesic which spirals to the same closed geodesic as γ. If J is not contained in K then γ is an isolated point of G_{one}, this is impossible because of Theorem 4.2. So $J \subset K$ and by maximality of J, $J = K$. □

4.2 Results

From the discussion in last section and Birman–Series Theorem, we get the following corollary.

Corollary 4.2. *The area of a cusp region is equal to the area of the union of of all maximal gaps.*

Hence we can prove McShane's identity.

Theorem 4.3. *On a once-punctured torus $\mathcal{S}_{1,1}$ with complete hyperbolic structure,*

$$\sum_{\gamma} \frac{1}{1 + \exp|\gamma|} = \frac{1}{2},$$

where the sum takes over all closed simple geodesics γ on $\mathcal{S}_{1,1}$ and $|\gamma|$ denotes the length of γ.

Proof. It is a direct corollary of Corollary 4.2 and Theorem 3.1. □

References

[1] J. S. Birman and C. Series. Geodesics with bounded intersection number on surfaces are sparsely distributed. *Topology*, 24(2):217–225, 1985.

[2] R. D. Canary, D. B. Epstein, and P. Green. *Notes on notes of Thurston*. University of Warwick Warwick, 1986.

[3] G. McShane. *A remarkable identity for lengths of curves*. PhD thesis, University of Warwick, 1991.

[4] G. McShane. Simple geodesics and a series constant over teichmuller space. *Inventiones mathematicae*, 132(3):607–632, 1998.

[5] K. Redzikultsava. Applied/plasma post seminar notes. URL https://www.overleaf.com/latex/templates/applied-slash-plasma-post-seminar-notes/vkxprmppmsjw.