May 26, 2008

THE WEIL-PETERSSON DISTANCE BETWEEN FINITE DEGREE COVERS OF PUNCTURED RIEMANN SURFACES AND RANDOM IDEAL TRIANGULATIONS

JEREMY KAHN AND VLADIMIR MARKOVIC

ABSTRACT. Let S and R be two hyperbolic finite area surfaces with cusps. We show that for every $\epsilon > 0$ there are finite degree unbranched covers $S_\epsilon \to S$ and $R_\epsilon \to R$, such that the Weil-Petersson distance between S_ϵ and R_ϵ is less than ϵ in the corresponding Moduli space.

1. INTRODUCTION

We say that a hyperbolic Riemann surface is of finite type if it has the finite area with respect to the underlining hyperbolic metric. Such surfaces are either closed or are obtained from closed surfaces after removing at most finitely punctures. All Riemann surfaces in this paper are hyperbolic and of finite type (except the unit disc/upper half space which is the universal cover of such surfaces). Let S and R be two finite type Riemann surfaces that are both either closed or both have at least one puncture. Then there is always a common holomorphic (possibly branched) cover of S and R. However a generic pair of such surfaces will not have a common holomorphic, unbranched finite degree cover. Except the universal cover, from now on all covers in this paper will be assumed to be holomorphic, unbranched and finite degree, so every time we use the term cover this will be understood. Since we assume that both S and R are either closed or have at least one puncture, one can find covers S_1 and R_1, of S and R respectively, that are quasiconformally equivalent (there is a quasiconformal map between them). This is equivalent to saying that S_1 and R_1 have the same genus $\mathbf{g}$ and the same number of punctures $\mathbf{n}$ (both S_1 and R_1 are of the type $(\mathbf{g}, \mathbf{n})$).

The well-known Ehrenpreis conjecture asserts that for a given $\epsilon > 0$ one can find covers S_1 and R_1, of S and R respectively, so that S_1 and R_1 are quasiconformally equivalent and the distance between them is less than ϵ. Since S_1 and R_1 are quasiconformally equivalent they belong to the same moduli space $\mathbf{M}_{g,n}$, where $\mathbf{g}$ is the genus and $\mathbf{n}$ is the number of punctures in S_1. Recall that $\mathbf{M}_{g,n}$ is the space of all hyperbolic metrics one can put on a surface that has genus $\mathbf{g}$ and n punctures. The distance between them is measured in terms of a metric that exists on $\mathbf{M}_{g,n}$. Originally the problem was posed in terms of any natural metric on the corresponding $\mathbf{M}_{g,n}$ (see [4]). Note that there are two cases of this conjecture, the first is when S and R have punctures, and the second when they are both closed.

Remark. In fact there is no easy way showing that one can find the corresponding covers S_1 and R_1 so that the distance between them in their moduli space is strictly less than the distance between S and R in their moduli space (providing that S and R are homeomorphic).

2000 *Mathematics Subject Classification.* Primary 20H10.

Recall that the Teichmüller space is the universal holomorphic cover of $\mathbf{M}_{g,n}$. The two standard (and most studied) metrics on $\mathbf{T}_{g,n}$ that are well defined on $\mathbf{M}_{g,n}$ are the Teichmüller metric and the Weil-Petersson metric. In this paper we prove this conjecture in the case of punctured surfaces and where the distance is measured with respect to the normalised Weil-Petersson metric. We stress that our results do not imply the case when the distance is measured with respect to the Teichmüller metric.

Remark. Let M be a finite type surface and $\pi : M_1 \to M$ be a cover (the genera of M and M_1 is $\mathbf{g}$ and $\mathbf{g}_1$ respectively and the number of punctures is $\mathbf{n}$ and $\mathbf{n}_1$ respectively). Then the covering map π induces an embedding $\iota : \mathbf{T}_{g,n} \to \mathbf{T}_{\mathbf{g}_1,\mathbf{n}_1}$. This embedding is an isometry of $\mathbf{T}_{g,n}$ onto its image with respect to the Teichmüller metric. However, if one takes the traditional definition of the Weil-Petersson metric this embedding increases the distance by multiplying it by the square root of the degree of the cover. Therefore it is important that we normalise the Weil-Petersson metric on $\mathbf{T}_{g,n}$ (and therefore on $\mathbf{M}_{g,n}$) by dividing it by the square root of the hyperbolic area of a surface of the type $(\mathbf{g},\mathbf{n})$ (every such surface has the area equal to $2\pi(2\mathbf{g}-2+\mathbf{n})$). After this normalisation the embedding ι becomes an isometry with respect to the Weil-Petersson metric as well.

Theorem 1.1. *Let S and R be two finite type Riemann surfaces that both have at least one puncture. Then given $\epsilon > 0$ one can find covers S_ϵ and R_ϵ of S and R respectively, so that S_ϵ and R_ϵ are quasiconformally equivalent and the Weil-Petersson distance between them is less than ϵ.*

We construct the covers S_ϵ and R_ϵ explicitely. We believe that it can be recovered from our construction that the degree of the covers $S_\epsilon \to S$ and $R_\epsilon \to R$ is of the order $P(\frac{1}{\epsilon})$, where P is a polynomial that depends on S and R. We believe that the degree of this polynomial is independent of S and R.

It has been shown in [6] that for every $\epsilon > 0$ there are covers of the Modular torus (this is the punctured torus that is isomorphic to $\mathbf{H}/G$ where G is a finite index subgroup of $\mathbf{PSL}(2,\mathbf{Z})$) that are not conformally the same but are ϵ close in the corresponding moduli space and with respect to the Teichmüller metric. See [2] for equivalent formulations of the Ehrenpreis conjecture (see also [8], [7] for related results).

We say that a Riemann surface S_0 is modular if S_0 is isomorphic to $\mathbf{H}/G$ where G is a finite index subgroup of $\mathbf{PSL}(2,\mathbf{Z})$. Such surfaces are characterised by having an ideal triangulation where all the sheears are equal to zero (a sheear coordinate that corresponds to an edge λ_i of an ideal triangulation is the signed hyperbolic distance between the normal projections to λ_i of the centres of the two ideal triangles that contain λ_i as their edge). Give a Riemann surface S, for every $r >> 0$ we construct a finite degree cover $S(r) \to S$ such that $S(r)$ has an ideal triangulation where the sheear coordinates are "small". Then there exists a modular surface $S_0(r)$ such that all the sheear coordinates of the corresponding ideal triangulation are equal to zero. One has to make this precise and in particular be able to estimate the Weil-Petersson distance between $S(r)$ and $S_0(r)$. We show that the Weil-Petersson distance between $S(r)$ and $S_0(r)$ less than $e^{-\frac{r}{8}}$. It can be recovered from the construction that the degree of the cover $S(r) \to S$ is less than $P(e^r)$, where $P(r)$ is a polynominal in r.

In Section 2. we discuss the Weil-Petersson distance and obtain the needed estimates of this distance in terms of the sheear coordinates. In Section 3. we develop the method of construction finite degree covers of S by gluing ideal immersed ideal triangles in S. In Section 4. we discuss measures on triangles and the notion of transport of measure. We state Theorem 4.1 which claims existence of certain measures on the space of immersed ideal triangles in S. We prove Theorem 1.1 using Theorem 4.1. In Section 5. we construct the measures from Theorem 4.1 and prove this theorem. Heavier computations needed in the proof of Theorem 4.1 are done in Section 6 and Section 7. In the appendix we prove an ergodic type theorem about the geodesic flow on a finite area hyperbolic surface with cusps. This theorem is most likely known but in the absence of an appropriate reference we offer a proof.

2. The shear coordinates and the Weil-Petersson metric

2.1. The Weil-Petersson metric. Let S be a Riemann surface of the type $(\mathbf{g},\mathbf{n})$. Let $\rho(z)|dz|$ be the line element for the hyperbolic metric on S (here $z = x+iy$ is the local parameter). Denote by $Q(S)$ the Banach space of holomorphic quadratic differentials on S with the norm given by

$$||\phi||_2 = \sqrt{\frac{1}{2\pi(2\mathbf{g}-2+\mathbf{n})}\int\limits_S \rho^{-2}(z)|\phi(z)|^2\, dxdy}.$$

Here $\text{Area}(S) = 2\pi(2\mathbf{g}-2+\mathbf{n})$ represents the hyperbolic area of S. Note that if the $\mathbf{n}\geq 1$ than elements of $Q(S)$ have at most first order poles at the punctures.

By $L^\infty(S)$ we denote the Banach space of Beltrami differentials on S. These are measurable $(-1,1)$ forms with the finite supremum norm $||\mu||_\infty$ for $\mu\in L^\infty(S)$. We introduce the equivalence relation on $L^\infty(S)$ by saying that $\mu\sim\nu$ if

$$\int\limits_S \mu\phi(z)\, dxdy = \int\limits_S \nu\phi(z)\, dxdy,$$

for every $\phi\in Q(S)$. The equivalence class of $\mu\in L^\infty(S)$ is denoted by $[\mu]$. The space $L^\infty(S)/\sim$ is a finite dimensional vector space. The induced supremum norm on $L^\infty(S)/\sim$ is given by

$$||[\mu]||_\infty = \inf_{\nu\in[\mu]} ||\nu||_\infty.$$

The Teichmüller space $\mathbf{T}_{g,n}$ is a complex manifold and its tangent space $T_S(\mathbf{T}_{g,n})$ at S is identified with the vector space $L^\infty/\sim$. The corresponding cotangent space $T^*_S(\mathbf{T}_{g,n})$ is identified with $Q(S)$. The Weil-Petersson pairing on $Q(S)$ is given by

$$\langle\phi,\psi\rangle_{WP} = \frac{1}{2\pi(2\mathbf{g}-2+\mathbf{n})}\int\limits_S \rho^{-2}(z)\phi(z)\overline{\psi(z)}\, dxdy.$$

The induced scalar product on $T_S(\mathbf{T}_{g,n}) = L^\infty(S)/\sim$ is called the Weil-Petersson product. The corresponding norm of a vector $[\mu]\in L^\infty(S)/\sim$ is given by

$$||[\mu]||_{WP} = \sup_{\phi\in Q(S)} \frac{1}{2\pi(2\mathbf{g}-2+\mathbf{n})||\phi||_2}\left|\int\limits_S \mu\phi\, dxdy\right|$$

Below we show how to define the norm $||[\mu]||_{WP}$ in terms of harmonic Beltrami differentials. The induced Riemannian metric on $\mathbf{T}_{g,n}$ is called the Weil-Petersson metric, and by d_{WP} we denote the corresponding distance on $\mathbf{T}_{g,n}$.

This definition of the Weil-Petersson distance is a modification of the usual one. If $P, Q \in \mathbf{T}_{g,n}$ than the usual Weil-Petersson distance between P and Q is given by $\sqrt{2\pi(2\mathbf{g}-2+\mathbf{n})}d_{WP}(P,Q)$. Clearly we have just rescaled the distance by the factor $\sqrt{2\pi(2\mathbf{g}-2+\mathbf{n})}$. If S_1 is a Riemann surface of the type (g_1, n_1) that covers S and if $\iota : \mathbf{T}_{g,n} \to \mathbf{T}_{\mathbf{g_1},\mathbf{n_1}}$ is the induced embedding, than ι is an isometric embedding with our definition of the Weil-Petersson distance. Also, according to [5] we have that $d_{WP}(P,Q)$ is strictly less than the Teichmüller distance between P and Q.

We have

Lemma 2.1. *Let $\mu \in L^{\infty}(S)$. Then*

$$||[\mu]||^2_{WP} \le 9||[\mu]||_{\infty} \left(\inf_{\nu \in [\mu]} \int_S |\nu| \rho^2(z)\, dxdy \right).$$

Proof. Let G be the covering group of Möbius transformations acting on the upper half plane $\mathbf{H}$ so that the Riemann surface S is isomorphic to $\mathbf{H}/G$. Let S_0 be a fundamental domain for the action of G. The lift of $\mu \in L^{\infty}(S)$ to $\mathbf{H}$ we also denote by μ. Recall that the density of the hyperbolic metric on $\mathbf{H}$ is given by $\rho^2(z) = y^{-2}$ where $z = x+iy \in \mathbf{H}$. Let $w = u+iv \in \mathbf{H}$ denote another complex parameter. Set

$$K(z,w) = \frac{1}{(\bar{z}-w)^4},$$

where $z, w \in \mathbf{H}$. The function $K(z,w)$ is the Bergman kernel for $\mathbf{H}$. The following are the well known properties of $K(z,w)$. We have

- $|K(z,w)| = |K(w,z)|$.
- For any Möbius transformation $g : \mathbf{H} \to \mathbf{H}$ we have

$$|K(g(z), g(w))||g'(z)|^2|g'(w)|^2 = |K(z,w)|, \tag{1}$$

- For every $z, w \in \mathbf{H}$ we have

$$\frac{4v^2}{\pi} \int_{\mathbf{H}} |K(z,w)|\, dxdy = \frac{4y^2}{\pi} \int_{\mathbf{H}} |K(z,w)|\, dudv = 1. \tag{2}$$

Let

$$\phi[\mu](w) = \frac{-12}{\pi} \int_{\mathbf{H}} \overline{\mu(z)} K(z,w)\, dxdy.$$

The differential $v^2\phi[\mu](w)$ is called the harmonic Beltrami differential. The Weil-Petersson norm $||[\mu]||_{WP}$ can be expressed as (see [1])

$$||[\mu]||^2_{WP} = \int_{S_0} v^2 |\phi[\mu](w)|^2\, dudv.$$

Let $\nu, \nu_1 \in [\mu]$ be any Beltrami dilatations from $[\mu]$ and use the same notation for the lifts of ν and ν_1 to $\mathbf{H}$. Then $\phi[\mu](w) = \phi[\nu](w) = \phi[\nu_1](w)$. We have

$$||[\mu]||^2_{WP} \leq \sup_{w \in \mathbf{H}} |v^2 \phi[\nu_1](w)| \left(\int\limits_{S_0} |\phi[\nu](w)| \, dudv \right).$$

From (2) we have

$$||[\mu]||^2_{WP} \leq 3||[\mu]||_\infty \int\limits_{S_0} |\phi[\nu](w)| \, dudv. \tag{3}$$

We have

$$\int\limits_{S_0} |\phi[\nu](w)| \, dudv = \frac{12}{\pi} \int\limits_{S_0} \left| \int\limits_{\mathbf{H}} \overline{\nu(z)} K(z,w) \, dxdy \right| dudv \leq$$

$$\leq \frac{12}{\pi} \int\limits_{S_0} \left(\int\limits_{\mathbf{H}} |\nu(z)||K(z,w)| \, dxdy \right) dudv.$$

We partition $\mathbf{H}$ into the sets $g(S_0)$, $g \in G$. This gives

$$\int\limits_{S_0} |\phi[\nu](w)| \, dudv \leq \frac{12}{\pi} \sum_{g \in G} \int\limits_{S_0} \left(\int\limits_{g(S_0)} |\nu(z)||K(z,w)| \, dxdy \right) dudv.$$

We have

$$\frac{12}{\pi} \sum_{g \in G} \int\limits_{S_0} \left(\int\limits_{g(S_0)} |\nu(z)||K(z,w)| \, dxdy \right) dudv =$$

$$= \frac{12}{\pi} \sum_{g \in G} \int\limits_{S_0} \left(\int\limits_{S_0} |\nu(g(z))||K(g(z),w)||g'(z)|^2 \, dxdy \right) dudv =$$

$$= \frac{12}{\pi} \sum_{g \in G} \int\limits_{g^{-1}(S_0)} \left(\int\limits_{S_0} |\nu(g(z))||K(g(z),g(w))||g'(z)|^2 |g'(w)|^2 \, dxdy \right) dudv.$$

Since $|\nu(g(z))| = |\nu(z)|$ and from (1) we get

$$\int\limits_{S_0} |\phi[\nu](w)| \, dudv \leq \frac{12}{\pi} \sum_{g \in G} \int\limits_{g^{-1}(S_0)} \left(\int\limits_{S_0} |\nu(z)||K(z,w)| \, dxdy \right) dudv =$$

$$= \frac{12}{\pi} \int\limits_{\mathbf{H}} \left(\int\limits_{S_0} |\nu(z)||K(z,w)| \, dxdy \right) dudv.$$

We exchange the integrals in the above inequality to get

$$\int\limits_{S_0} |\phi[\nu](w)|\, dudv \leq \frac{12}{\pi} \int\limits_{S_0} \left(\int\limits_{\mathbf{H}} |\nu(z)||K(z,w)|\, dudv \right) dxdy.$$

From (2) we have that

$$\frac{12}{\pi} \int\limits_{\mathbf{H}} |K(z,w)|\, dudv \leq 3y^{-2},$$

which shows that

$$\int\limits_{S_0} |\phi[\nu](w)|\, dudv \leq 3 \int\limits_{S_0} y^{-2}|\nu(z)|\, dxdy = 3 \int\limits_{S} \rho^2(z)|\nu(z)|\, dxdy,$$

where $\rho^2(z)$ is the density of the hyperbolic metric on S. Together with (3) this proves the lemma.

□

2.2. **The Shear coordinates for $\mathbf{T}_{g,n}$.** The notation we introduce here remains valid throughout this section. Fix a surface S of genus $\mathbf{g}$ and with $\mathbf{n} \geq 1$ punctures (here S is not assumed to be a Riemann surface, in fact we will equip S with various complex structures). By $\text{Cusp}(S) = \{c_1(S), c_2(S), ..., c_{\mathbf{n}}(S)\}$ we denote the set of punctures (recall that $\overline{S}$ is a closed surface and S is obtained by removing the set $\text{Cusp}(S)$ from $\overline{S}$). Let τ be an ideal triangulation of S. This means that τ is a triangulation of $\overline{S}$ where the vertex set is exactly equal to $\text{Cusp}(S)$ (from now on all triangulations will be ideal triangulations and we will not use the term ideal anymore). By $\lambda = \{\lambda_1, ..., \lambda_{|\lambda|}\}$ we denote the ordered set of edges of the triangles from τ. Here $|\lambda|$ denotes the total number of edges. For topological reasons we have $|\lambda| = 6\mathbf{g} - 6 + 3\mathbf{n}$. Let $|\tau|$ denote the total number of triangles in τ. We have $|\tau| = 2(2\mathbf{g} - 2 + \mathbf{n})$ and therefore the equality

$$\frac{2}{3}|\lambda| = |\tau|,$$

holds. We say that two triangulations τ and τ_1 of S are isotopic if there is a homeomorphism $f : \overline{S} \to \overline{S}$ that pointwise preserves the set $\text{Cusp}(S)$ and that is homotopic to the identity map on S modulo the set $\text{Cusp}(S)$.

Fix a triangulation τ of S. We define the set $X(\tau) \subset \mathbf{R}^{|\lambda|}$ as follows. Let $c_i(S) \in \text{Cusp}(S)$ and let $\lambda_{i(1)}, .., \lambda_{i(k)}$ be the subset of edges from the set λ that have $c_i(S)$ as an end point. Also, let $\sigma : \lambda \to \{1, 2\}$ be defined so that for $\lambda_i \in \lambda$ we have $\sigma(\lambda_i) = 1$ if the endpoints of λ_i represent different punctures on S and $\sigma(\lambda_i) = 2$ if the two endpoints represent the same puncture on S. Let $\mathbf{r} = (\mathbf{r}_1, ..., \mathbf{r}_{|\lambda|}) \in \mathbf{R}^{|\lambda|}$. Then $\mathbf{r} \in X(\tau)$ if for every $i = 1, ..., \mathbf{n}$ we have

$$\sum_{j=1}^{j=k} \sigma(\lambda_{i(j)})\mathbf{r}_{i(j)} = 0.$$

Clearly $X(\tau)$ is a linear subset of $\mathbf{R}^{|\lambda|}$. In particular, we have $0 = (0, ..., 0) \in X(\tau)$.

Now we define the map $F_\tau : X(\tau) \to \mathbf{T}_{g,n}$. Recall that $\mathbf{T}_{g,n}$ is the space of complete hyperbolic metrics on S with marking. To a point $\mathbf{r} = (\mathbf{r}_1, ...\mathbf{r}_{|\lambda|}) \in X(\tau)$

we associate an element $F_\tau(\mathbf{r}) \in \mathbf{T}_{g,n}$ that is represented by the marked hyperbolic metric with the following properties. There exists a triangulation τ' (with the corresponding set of edges λ') of S that is isotopic to τ and such that all the edges in λ' are geodesics with respect to this metric and so that the following holds (a triangulation will be called a geodesic triangulation if all the edges are geodesics in the corresponding hyperbolic metric). Let $\lambda'_i \in \lambda'$ and let T_1 and T_2 be the two triangles from the triangulation τ' that have λ'_i as an edge. We lift λ'_i, T_1 and T_2 to the universal cover $\mathbf{H}$ (the upper half space). The edge λ'_i lifts to the geodesic that connects 0 and ∞. One of the triangles T_1 or T_2 (depending on how we lift λ'_i) lifts to the ideal triangle with vertices at $-1, 0, \infty$ and the other triangle lifts to the triangle with vertices $0, e^{\mathbf{r}_i}, \infty$. This requirement defines the map F_τ. This definition does not depend on how we lift the geodesic λ'_i (there are exactly two ways we can lift it). The number $\mathbf{r}_i$ can also be defined as the signed hyperbolic distance between the points on the geodesic λ'_i that are the orthogonal projections of the centres of the triangles of T_1 and T_2 respectively (equivalently one can see these two points as the feet of the perpendiculars from the two vertices of T_1 and T_2 respectively that are opposite to λ'_i).

This maps was originally defined by Thurston (this concept has been developed by Bonahon, Penner, Fock, Chekhov...). The corresponding coordinates on $\mathbf{T}_{g,n}$ are called the shear coordinates.

Proposition 2.1. *The map $F_\tau : X(\tau) \to \mathbf{T}_{g,n}$ is well defined real analytic homeomorphism.*

See [3] for the proof of this theorem. This parametrisation of $\mathbf{T}_{g,n}$ depends on the choice of the triangulation τ and different triangulations produce different parametrisation.

Recall that the Farey tessellation $\mathcal{F}$ is the ideal triangulation of $\mathbf{H}$ with the property that for any two adjacent triangles the feet of the perpendiculars that are dropped from the vertices of these two triangles, that are opposite to their common edge, coincide. To make $\mathcal{F}$ be a unique tessellation satisfying this property we require that one its triangles is the one with the vertices at $0, 1, \infty$. It is well known that this triangulation is preserved by the action of the group $\mathbf{PSL}(2, \mathbf{Z})$.

Proposition 2.2. *Let τ be a triangulation of a finite type surface S (S has at least one puncture). Then the Riemann surface that corresponds to $F_\tau(0)$ is obtained as the quotient of $\mathbf{H}$ by a finite index subgroup of $\mathbf{PSL}(2, \mathbf{Z})$.*

Proof. Let R be the Riemann surface that corresponds to $F_\tau(0)$. We lift the corresponding geodesic triangulation τ' of R to $\mathbf{H}$ and assume that the triangle with vertices at $0, 1, \infty$ belongs to this lift. Let G be the corresponding covering group of Möbius transformation acting on $\mathbf{H}$. Since R corresponds to $F_\tau(0)$ we conclude that the corresponding tessellation of $\mathbf{H}$ is the Farey tessellation $\mathcal{F}$ and G preserves $\mathcal{F}$. On the other hand, every Möbius transformation that preserves $\mathcal{F}$ must be in $\mathbf{PSL}(2, \mathbf{Z})$. This shows that G is a subgroup of $\mathbf{PSL}(2, \mathbf{Z})$. The fact that G is of finite index follows from the assumption that S is of finite type so R has finite hyperbolic area. □

This simple proposition is important for us.

For $\mathbf{r} \in X(\tau)$ we define the supremum norm $||\mathbf{r}||_\infty = \max\{\mathbf{r}_1, ..., \mathbf{r}_{|\lambda|}\}$ as usual (this norm does not depend on τ of course). Next, we define a norm on $X(\tau)$

that depends on τ. We define the Oscillation norm $O_\tau(\mathbf{r})$ for $\mathbf{r} \in X(\tau)$ as follows. Let G be the covering group of Möbius transformations acting on $\mathbf{H}$ so that S is isomorphic to $\mathbf{H}/G$ (here S has the complex structure that corresponds to $F_\tau(\mathbf{r})$). Let $\{\lambda_{i(1)}, ..., \lambda_{i(k)}\}$ be a k-tuple of edges from λ. We say that this k-tuple is a k-tuple of consecutive edges if we can find the lifts $\{\lambda'_{i(1)}, ..., \lambda'_{i(k)}\}$ to $\mathbf{H}$ so that each curve $\lambda'_{i(j)}$ has ∞ as its endpoint, and so that each curve $\lambda'_{i(j)}$ is to the left of the curve $\lambda'_{i(j+1)}$, where $j = 1, ...k_1$. In this case we also say that the k-tuple of consecutive edges $\{\lambda_{i(1)}, ..., \lambda_{i(k)}\}$ is left oriented (one similarly defines a right oriented k-tuple of consecutive edges).

Set

$$O_\tau(\mathbf{r}) = \sup_{\{\lambda_{i(1)},...,\lambda_{i(k)}\}} |\mathbf{r}_{i(1)} + ... + \mathbf{r}_{i(k)}|,$$

where the supremum is taken among all consecutive k-tuples $\{\lambda_{i(1)}, ..., \lambda_{i(k)}\}$ (and for any $k \in \mathbf{N}$). Note that if k is equal to the number of all edges that enter the puncture that corresponds to ∞ (an edge is counted twice if the puncture on S that corresponds to ∞ is equal to both of its endpoints) then by definition of $X(\tau)$ we have $\mathbf{r}_{i(1)} + ... + \mathbf{r}_{i(k)} = 1$. This shows that the supremum in the above definition is achieved and this shows that $O_\tau(\mathbf{r})$ is a well defined non-negative real number. Note that $||\mathbf{r}||_\infty \leq O_\tau(\mathbf{r})$.

2.3. **Estimating the Weil-Petersson distance in terms of the shear coordinates.** Our aim here is to estimate from above the Weil-Petersson distance $d_{WP}(F_\tau(0), F_\tau(\mathbf{r}))$ for a given $\mathbf{r} \in X(\tau)$. We make this estimate in terms of the vector $\mathbf{r}$ (and under certain assumptions on $\mathbf{r}$). Until the end of this section $\mathbf{r} \in X(\tau)$ is a fixed vector.

Let $\psi : [0, 1] \to \mathbf{T}_{g,n}$ be given by $\psi(t) = F_\tau(t\mathbf{r})$. The map ψ is differentiable (since F_τ is differentiable), and we compute its first derivative in order to estimate the distance. For $t \in [0, 1]$ let S_t be the Riemann surface that corresponds to the point $F_\tau(t\mathbf{r}) \in \mathbf{T}_{g,n}$. Fix $t_0 \in [0, 1]$ and identify $\mathbf{T}_{g,n}$ with $\text{Teich}(S_{t_0})$ in the standard way. In order to estimate the distance $d_{WP}(F_\tau(0), F_\tau(\mathbf{r}))$ we estimate the Weil-Petersson norm of the vector $\frac{\partial\psi}{\partial t}(t_0)$ in the tangent space of $\mathbf{T}_{g,n}$ at the point $F_\tau(t_0\mathbf{r})$. First we construct an explicit quasiconformal map $f_t : S_{t_0} \to S_t$ so that the pair (S_t, f_t) represents the point $\psi(t)$ in $\text{Teich}(S_{t_0})$ (the requirement is that f_t is homotopic to the identity as a map of S onto itself).

Remark. The homotopy class of map f_t has been studied by Penner-Saric in [9]. Passing to the universal cover of S_t, they explicitely construct the quasisymmetric map of the unit circle that determines the homotopy class of f_t in terms of the corresponding ideal triangulation of the unit disc. This has various important applications.

Let $\tau(t)$ be the geodesic triangulation of S_t that is homotopic to τ. By $\lambda(t)$ we denote the corresponding set of edges (these edges are now geodesics). For a triangle $T \in \tau$ we denote by $T(t)$ the corresponding triangle in $\tau(t)$ (for $\lambda_i \in \lambda$ we denote by $\lambda_i(t)$ the corresponding element of $\lambda(t)$). Let $\text{ct}(T(t))$ be the geometric centre of $T(t)$. For adjacent triangles $T_1, T_2 \in \tau$ let $l(T_1(t), T_2(t))$ denote the geodesic segment between the centres $\text{ct}(T_1(t))$ and $\text{ct}(T_2(t))$. We define the map f_t at the centres $\text{ct}(T(t_0))$ by setting $f_t(\text{ct}(T(t_0))) = \text{ct}(T(t))$. We define f_t on each

$l(T_1(t_0), T_2(t_0))$ so that $f_t(l(T_1(t_0), T_2(t_0))) = l(T_1(t), T_2(t))$ and by the requirement that f_t stretches the hyperbolic distance by the factor

$$\frac{|l(T_1(t), T_2(t))|}{|l(T_1(t_0), T_2(t_0))|},$$

where $|l(T_1(t), T_2(t))|$ and $|l(T_1(t_0), T_2(t_0))|$ are the corresponding hyperbolic lengths. If $\lambda_i(t) \in \lambda$ is the common edge for T_1 and T_2 then we have

$$f_t(l(T_1(t_0), T_2(t_0)) \cap \lambda_i(t_0)) = l(T_1(t), T_2(t)) \cap \lambda_i(t),$$

because the point $l(T_1(t), T_2(t)) \cap \lambda_i(t)$ is always the middle point (in terms of the hyperbolic distance) of the geodesic segment $l(T_1(t), T_2(t))$.

Fix $T \in \tau$ and let $T_1, T_2, T_3 \in \tau$ be triangles adjacent to T. The triangle $T(t)$ is partitioned into sets $E_1(t), E_2(t), E_3(t)$ where these three sets are separated by the segments $l(T(t), T_i(t)) \cap T(t)$, $i = 1, 2, 3$. In particular, the boundary of set $E_1(t)$ is the union of the curve $(l(T(t), T_1(t)) \cap T(t)) \cup (l(T(t), T_2(t)) \cap T(t))$ and the two geodesic rays lying on the edges that separate the pairs $T(t), T_1(t)$ and $T(t), T_2(t)$ respectively. Denote by $\lambda_{i(1)}(t)$ and $\lambda_{i(2)}(t)$ the corresponding edges that separate the pairs $T(t), T_1(t)$ and $T(t), T_2(t)$ respectively. The map f_t is already defined on the part of the boundary of $E_1(t_0)$ and we define it on the rest of $E_1(t_0)$ as follows.

We use the same notation for the lifts of triangles from $\tau(t)$ and the edges from $\lambda(t)$ to the universal cover $\mathbf{H}$. We assume that for every $t \in [0, 1]$ the triangle $T(t) \in \tau(t)$ lifts to the triangle in $\mathbf{H}$ that has the vertices at $0, 1, \infty$ (our definition of f_t does not depend on this normalisation). We may assume that $T_1(t)$ is to the left of the triangle $T(t)$ in the universal cover. Then the edges $\lambda_{i(1)}$ and $\lambda_{i(2)}$ lift to the geodesics in $\mathbf{H}$ with the endpoints $0, \infty$ and $1, \infty$ respectively. We have already seen that the two vertices of the triangle $T_1(t)$ are 0 and ∞. The third vertex of $T_1(t)$ is at the point $-e^{-t\mathbf{r}_{i(1)}}$ (see Figure 1). Similarly, the vertices of $T_2(t)$ are at $1, (1 + e^{t\mathbf{r}_{i(2)}}), \infty$. We need to keep track of the vertices for $T_1(t)$ and $T_2(t)$ in order to be able to verify later that the map f_t is well defined on the geodesics $\lambda_{i(1)}(t_0)$ and $\lambda_{i(2)}(t_0)$.

Set

$$L_t = (l(T(t), T_1(t)) \cap T(t)) \cup (l(T(t), T_2(t)) \cap T(t)).$$

Let $u(t, .) : [0, 1] \to (0, \infty)$ so that L_t is the graph of the function $u(t, x)$. The function $u(t, x)$ has all derivatives everywhere except at the point $\frac{1}{2}$ (but it has both left and the right derivatives at this point). At the point t_0 we have that the function $u(t_0, x)$ depends only on the value of the shear coordinates $\mathbf{r}_{i(1)}$ and $\mathbf{r}_{i(2)}$. If we fix an upper bound on the sum $|\mathbf{r}_{i(1)}| + |\mathbf{r}_{i(2)}|$ we have that the set of such functions $u(t, .) : [0, 1] \to (0, \infty)$ is compact in the C^∞ topology on both intervals $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$. This shows that there is a constant $C_1 > 0$ that depends only on $|\mathbf{r}_{i(1)}| + |\mathbf{r}_{i(2)}|$ so that

$$||u(t_0, .)||_\infty||, ||\frac{1}{u(t_0, .)}||_\infty, ||u_x(t_0, .)||_\infty \leq C_1. \tag{4}$$

The set $E_1(t)$ is given by

$$E_1(t) = \{z \in \mathbf{H} : 0 \leq x \leq 1, y \geq u(t, x)\}.$$

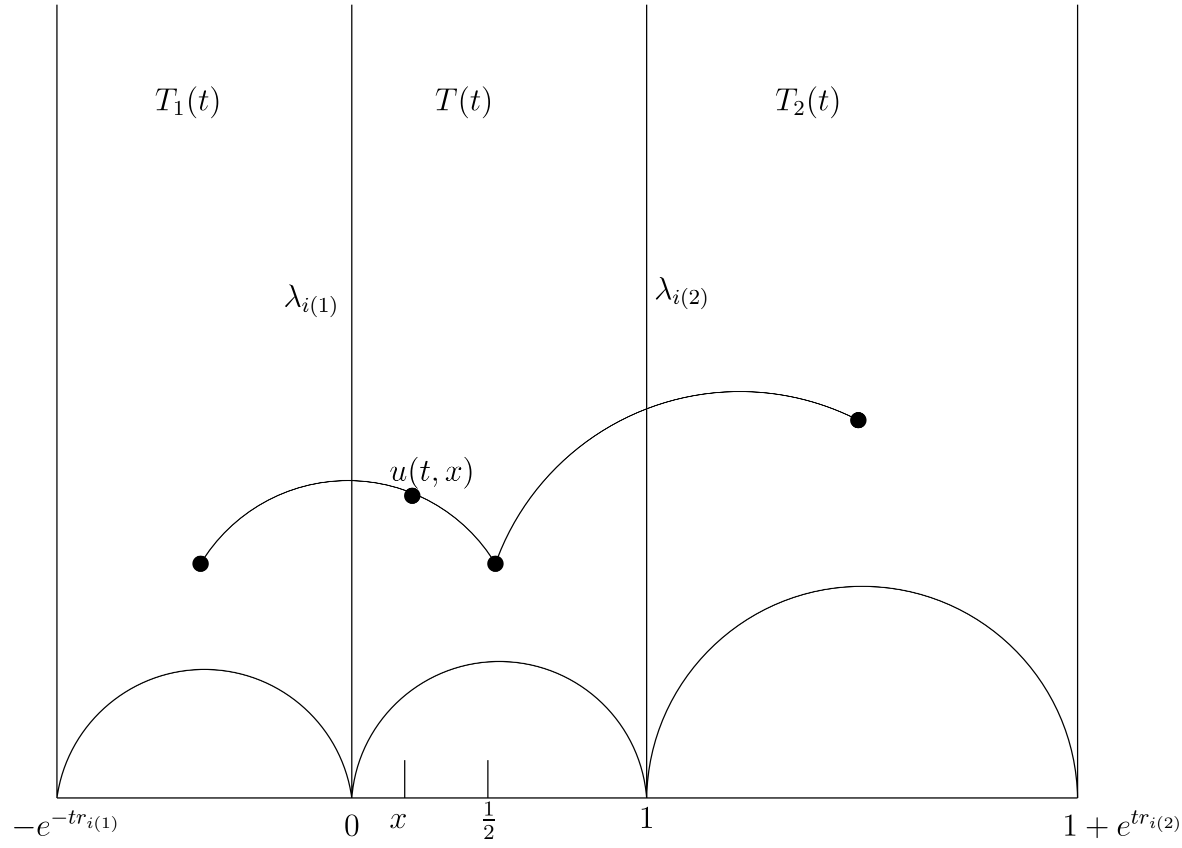

FIGURE 1

The map f_t is already defined on L_{t_0}. Let $\alpha(t,.):[0,1]\to[0,1]$ and $\beta(t,.):[0,1]\to(0,\infty)$ be the functions so that

$$f_t(x+iu(t_0,x))=\alpha(t,x)+i\beta(t,x).$$

The functions $\alpha(t,.)$ and $\beta(t,.)$ depend only on the values $\mathbf{r}_{i(1)}$ and $\mathbf{r}_{i(2)}$ and both of them have all derivatives at every point in $(0,1)$ except $\frac{1}{2}$ (but it has the left and the right derivative at this point). Note that $\alpha(t,.)$ fixes the points $0,\frac{1}{2},1$. Also, $\alpha(t_0,x)=x$ and $\beta(t_0,x)=u(t_0,x)$. Again, by fixing an upper bound on the sum $|\mathbf{r}_{i(1)}|+|\mathbf{r}_{i(2)}|$ we have that the set of all such functions $\alpha(t,.):[0,1]\to[0,1]$ and $\beta(t,.):[0,1]\to(0,\infty)$ is compact in the C^{∞} topology on the intervals $[0,\frac{1}{2}]$ and $[\frac{1}{2},1]$. This shows that there are functions $\varphi_1,\varphi_2:[0,1]\setminus\frac{1}{2}\to\mathbf{R}$ that depend only on $|\mathbf{r}_{i(1)}|+|\mathbf{r}_{i(2)}|$ so that

(5)
$$\alpha_x(t,x)=1+\varphi_1(x)(t-t_0)+o(t-t_0),\ \beta_x(t,x)=u_x(t_0,x)+\varphi_2(x)(t-t_0)+o(t-t_0),$$

for every $x\in(0,1)$, $x\neq\frac{1}{2}$. Moreover, the functions $\alpha(t,x)$ and $\beta(t,x)$ depend on the real variables $\mathbf{r}_{i(1)}$ and $\mathbf{r}_{i(2)}$ and have all derivatives with respect to these variables too. If we fix an upper bound on the sum $|\mathbf{r}_{i(1)}|+|\mathbf{r}_{i(2)}|$ then the set of functions $\varphi_1(x)$ and $\varphi_2(x)$ is compact in the variables x, $\mathbf{r}_{i(1)}$ and $\mathbf{r}_{i(2)}$ in the C^∞ topology. In particular, the derivatives of φ_1 and φ_2 with respect to $\mathbf{r}_{i(1)}$ and $\mathbf{r}_{i(2)}$ are bounded from above, and this bound depends only on the sum $|\mathbf{r}_{i(1)}|+|\mathbf{r}_{i(2)}|$. This shows that there is a constant $C_2>0$ that depends only on the upper bound on the sum $|\mathbf{r}_{i(1)}|+|\mathbf{r}_{i(2)}|$, so that

$$||\varphi_1||_\infty, ||\varphi_2||_\infty \leq C_2(|\mathbf{r}_{i(1)}|+|\mathbf{r}_{i(2)}|). \tag{6}$$

The estimate (6) will be used in the case when the sum $|\mathbf{r}_{i(1)}|+|\mathbf{r}_{i(2)}|$ is small.

We define the map f_t so that $f_t(E_1(t_0))=E_1(t)$. We have

$$f_t(z)=\alpha(t,x)+i(\beta(t,x)+\gamma(t,z)),$$

where the function $\gamma: E_1(t_0)\to[0,\infty)$ is defined as follows. By $\lambda_{i(1)}(t_0)$, $\lambda_{i(2)}(t_0)$, ..., $\lambda_{i(k)}(t_0)$ we denote a k-tuple of consecutive edges (see the definition above), starting with the edge $\lambda_{i(1)}$. Let $z\in E_1(t_0)$. Then $y-u(t_0,x)\geq 0$. If $y-u(t_0,x)\geq(1-x)$ then let $k(z)\geq 2$ be the smallest integer so that

$$y-u(t_0,x)\leq(1-x)+\sum_{j=2}^{j=k(z)} e^{t_0(\mathbf{r}_{i(2)}+\ldots+\mathbf{r}_{i(j)})}.$$

In this case let $p(z)\in[0,1]$ be determined by the identity

$$y-u(t_0,x)=(1-x)+\sum_{j=2}^{j=k(z)-1} e^{t_0(\mathbf{r}_{i(2)}+\ldots+\mathbf{r}_{i(j)})}+p(z)e^{t_0(\mathbf{r}_{i(2)}+\ldots+\mathbf{r}_{i(k(z))})}. \tag{7}$$

If $0<y<1-x$ then $p(z)\in[0,1]$ is given by $y-u(t_0,x)=p(z)(1-x)$.

The function $p(z)$ is defined for $y\geq u(t_0,x)$ and $0\leq x\leq 1$ and it is continuous there, while it is differentiable outside a discrete set of smooth curves. From (7) for $y-u(t_0,x)\geq(1-x)$ we have

$$p_x(z)=(1-u_x(t_0,x))e^{-t_0(\mathbf{r}_{i(2)}+\ldots+\mathbf{r}_{i(k(z))})},\ p_y(z)=e^{-t_0(\mathbf{r}_{i(2)}+\ldots+\mathbf{r}_{i(k(z))})}, \tag{8}$$

where defined. For $y-u(t_0,x)\geq(1-x)$ set

$$\gamma(t,z)=(1-x)+\sum_{j=2}^{j=k(z)-1} e^{t(\mathbf{r}_{i(2)}+\ldots+\mathbf{r}_{i(j)})}+p(z)e^{t(\mathbf{r}_{i(2)}+\ldots+\mathbf{r}_{i(k(z))})}.$$

For $y-u(t_0,x)<1-x$ set

$$\gamma(t,z)=p(z)(1-x)=y-u(t_0,x).$$

For $y-u(t_0,x)\geq(1-x)$ it follows from (8) that

$$\gamma_x(t,z)=(1-u_x(t_0,x))e^{(t-t_0)(\mathbf{r}_{i(2)}+\ldots+\mathbf{r}_{i(k(z))})}-1,\ \gamma_y(t,z)=e^{(t-t_0)(\mathbf{r}_{i(2)}+\ldots+\mathbf{r}_{i(k(z))})}. \tag{9}$$

For $y-u(t_0,x)<1-x$ we have

$$\gamma_x(t,z)=-u_x(t_0,x),\ \gamma_y(t,z)=1. \tag{10}$$

As we stated above we define $f_t(z)$ as

$$f_t(z)=\alpha(t,x)+i(\beta(t,x)+\eta(t,z)).$$

Clearly $f_t: E_1(t_0)\to E_1(t)$ is a homeomorphism. It is differentiable outside a discrete set of smooth curves. By repeating the same process we define f_t on every triangle. Every triangles can mapped by a Möbius transformation to the triangle with vertices $0, 1, \infty$ and this how we defined the map f_t on every triangle. It is directly seen that the definition of f_t on an edge $\lambda_i(t_0)$ does not depend on the choice of the triangle that has $\lambda_i(t_0)$ in its boundary. For example consider the triangle $T_1(t_0)$ that is adjacent to $T(t_0)$ (recall that $T(t)$ has the vertices at 0, 1, ∞ and $T_1(t)$ has the vertices $-e^{-t\mathbf{r}_{i(1)}}$, 0, ∞). Let $A_t(w) = e^{t\mathbf{r}_{i(1)}} + 1$. Then $A_t(T_1(t)) = T(t)$. We define $g_t : T_1(t_0) \to T_1(t)$ by $(A_t)^{-1} \circ f_t \circ A_{t_0}$ where $f_t : T(t_0) \to T(t) = T(t_0)$. It is elementary to see that $f_t = g_t$ on $\lambda_{i(1)}(t_0)$. We define f_t on $T_1(t_0)$ by setting $f_t = g_t$.

The Beltrami dilatation of f_t is given by

$$(\frac{\partial f_t}{\partial \bar{z}} / \frac{\partial f_t}{\partial z})(t,z) = \frac{(\alpha_x(t,x) - \gamma_y(t,z)) + i(\beta_x(t,x) + \gamma_x(t,z))}{(\alpha_x(t,x) + \gamma_y(t,z)) + i(\beta_x(t,x) + \gamma_x(t,z))}.$$

Set

$$\mu(z) = \frac{\partial(\frac{\partial f_t}{\partial \bar{z}} / \frac{\partial f_t}{\partial z})}{\partial t}(t_0, z).$$

If $y - u(t_0,x) \geq 1 - x$ from (5) and (9) we compute

$$\mu(z) = \frac{1}{2}\big(\varphi_1(x) - (\mathbf{r}_{i(2)} + ... + \mathbf{r}_{i(k(z))})\big) +$$

(11)

$$+\frac{i}{2}\big(\varphi_2(x) + (1 - u_x(t_0,x))(\mathbf{r}_{i(2)} + ... + \mathbf{r}_{i(k(z))})\big).$$

If $y - u(t_0,x) < 1 - x$ then

$$\mu(z) = \frac{1}{2}(\varphi_1(x) + i\varphi_2(x)). \tag{12}$$

Going back to our path $\psi : [0,1] \to \mathbf{T}_{g,n}$ we have that $\mu \in L^\infty(S_{t_0})$ represents the tangent vector $\frac{\partial \psi}{\partial t}(t_0)$ in the tangent space of $\mathbf{T}_{g,n}$ at the point $F_\tau(t_0\mathbf{r})$.

Next, we derive two estimates needed in the proofs of Theorem 2.1 and Theorem 2.2 below. From (4), (6), (11), (12) we have

$$|\mu(z)| \leq C_2(|\mathbf{r}_{i(1)}| + |\mathbf{r}_{i(2)}|) + \frac{2 + C_1}{2} O_\tau(\mathbf{r}).$$

which shows that

$$||\mu||_\infty \leq 2||\mathbf{r}||_\infty C_2 + \frac{2 + C_1}{2} O_\tau(\mathbf{r}).$$

Set $C_3 = 2C_2 + \frac{2+C_1}{2}$. Since $||\mathbf{r}||_\infty \leq O_\tau(\mathbf{r})$ we get

$$||\mu||_\infty \leq C_3 O_\tau(\mathbf{r}). \tag{13}$$

Remark. It follows from (13) that the Teichmüller norm of the tangent vector $\frac{\partial \psi}{\partial t}(t_0)$ is less or equal to $C_3 O_\tau(\mathbf{r})$ for every $0 < t_0 < 1$. Let S_m, $m \in \mathbf{N}$, be a sequence of surfaces equipped with triangulations τ_m and let $\mathbf{r}_m \in X(\tau_m)$ be a sequence of vectors so that $O_{\tau_m}(\mathbf{r}_m) \to 0$, $m \to \infty$. Since $||\mathbf{r}_m||_\infty \leq O_{\tau_m}(\mathbf{r}_m)$ the constant C_3 in (13) does not depend on m. This implies that the Teichmüller distance between the points $F_{\tau_m}(0)$ and $F_{\tau_m}(\mathbf{r}_m)$ tends to 0 when $m \to \infty$.

Let $0 < \delta < 1$. Let $k \in \mathbf{N}$ such that $k \geq 2$ and such that $k\delta \leq 1$. Assume that for every $1 \leq j \leq k$ we have $|\mathbf{r}_{i(j)}| \leq \delta$. For $z \in E_1(t_0)$ which satisfies that

$$y - u(t_0, x) \leq (k-1)e^{-k\delta t_0},$$

from (7) we get that $k(z) \leq k$. Then from (11) we have

$$|\mu(z)| \leq \frac{1}{2}(2\delta C_2 + k\delta + 2\delta C_2 + k\delta + C_1 k\delta) \leq C_4 k\delta, \tag{14}$$

where C_4 is a constant that depends only on the upper bound for $|\mathbf{r}_{i(1)}| + |\mathbf{r}_{i(2)}| \leq 2\delta < 1$ that is C_4 is a universal constant.

Furthermore

$$\int\limits_{E_1(t_0)} |\mu(z)|\rho^2(z)\, dxdy = \int\limits_{E_1(t_0)} |\mu(z)| y^{-2}\, dxdy =$$

$$= \int\limits_0^1 \left(\int\limits_{u(t_0,x)}^{(k-1)e^{-k\delta t_0}} |\mu(z)| y^{-2}\, dy + \int\limits_{(k-1)e^{-k\delta t_0}}^{\infty} |\mu(z)| y^{-2}\, dy \right) dx.$$

In the first integral above we estimate $|\mu(z)|$ by (14) and in the second integral we estimate $|\mu(z)|$ by (13). Note that $0 \leq t_0 \leq 1$ and since $k \geq 2$ we have $\frac{1}{k-1} \leq \frac{2}{k}$. From the assumption $k\delta \leq 1$ we have $e^{k\delta} \leq e$, therefore

$$\int\limits_{(k-1)e^{-k\delta t_0}}^{\infty} |\mu(z)| y^{-2}\, dy \leq 10 \frac{O_\tau(\mathbf{r})}{k}.$$

We conclude

$$\int\limits_{E_1(t_0)} |\mu(z)|\rho^2(z)\, dxdy \leq C_5\big(k\delta + \frac{O_\tau(\mathbf{r})}{k}\big), \tag{15}$$

where C_5 depends only on the upper bound for $|\mathbf{r}_{i(1)}| + |\mathbf{r}_{i(2)}| \leq 2\delta < 1$, that is C_5 is a universal constant.

Theorem 2.1. *Let S be a surface of type $(\mathbf{g}, \mathbf{n})$ (where $n \geq 1$) and let τ be a triangulation on S (where λ denotes the corresponding set of edges). Fix $\mathbf{r} \in X(\tau)$. Let $0 < \delta < 1$ and denote by $N_\tau(\delta)$ the number of edges for which the absolute value of the corresponding shear coordinate is greater than δ. Then there is a constant C that depends only on $||\mathbf{r}||_\infty$ so that for any $k \in \mathbf{N}$ such that $k\delta \leq 1$ and $k \geq 2$ we have*

$$d_{WP}(F_\tau(0), F_\tau(\mathbf{r})) \leq C_7 \sqrt{O_\tau(\mathbf{r}) \left(k\delta + \frac{O_\tau(\mathbf{r})}{k} + \frac{N_\tau(\delta)}{|\lambda|} O_\tau(\mathbf{r}) \right)}. \tag{16}$$

Remark. Important point in this theorem is that the constant C does not depend on choice of the triangulation τ or on the type $(\mathbf{g}, \mathbf{n})$. It only depends on the upper bound on $||\mathbf{r}||_\infty$.

Proof. Let $\widetilde{\lambda}=\{\lambda_i \in \lambda : |\mathbf{r}_i|>\delta\}$. Then $|\widetilde{\lambda}|=N_{\tau}(\delta)$. We enlarge the set $\widetilde{\lambda}$ to the set $\widehat{\lambda}$ as follows. An edge from λ belongs to $\widehat{\lambda}$ if it belongs to a consecutive k-tuple of edges that starts at an edge from $\widetilde{\lambda}$ (here we take both the left and the right oriented consecutive k-tuples). Note that

$$|\widehat{\lambda}| \le 4k|\widetilde{\lambda}|.$$

By $\widehat{\tau}$ we denote the set of triangles that have at least one of its edges in $\widehat{\lambda}$. Note that

$$|\widehat{\tau}| \le 2|\widehat{\lambda}| \le 8k|\widetilde{\lambda}|. \tag{17}$$

To estimate $d_{WP}(F_{\tau}(o),F_{\tau}(\mathbf{r}))$ we need to estimate $||\frac{\partial\psi}{\partial t}(t_0)||_{WP}$ for every $t_0 \in [0,1]$. Here $\psi:[0,1]\to \mathbf{T}_{g,n}$ is the path defined above and we have

$$d_{WP}(F_{\tau}(0),F_{\tau}(\mathbf{r})) \le \int\limits_0^1 ||\frac{\partial\psi}{\partial t}(t)||_{WP}dt. \tag{18}$$

Assume that a triangle T does not belong to $\widehat{\tau}$ and let $\lambda_{i(1)}$ be its edge. As above, by $\lambda_{i(1)},...,\lambda_{i(k)}$, we denote the consecutive k-tuple of edges. For each $1\le j\le k$, we have $|\mathbf{r}_{i(j)}|\le\delta$. Applying (15) on $E_1(t_0)$ (and repeating the same on $E_2(t_0)$ and $E_3(t_0)$) we get

$$\int\limits_{T(t_0)} |\mu(z)|\rho^2(z)\,dxdy \le 3C_5\big(k\delta+\frac{O_{\tau}(\mathbf{r})}{k}\big).$$

On the other hand, for a triangle $T\in\widehat{\tau}$ from (13) we get the estimate

$$\int\limits_{T(t_0)} |\mu(z)|\rho^2(z)\,dxdy \le \pi C_3 O_{\tau}(\mathbf{r}).$$

This implies that (we express the hyperbolic area of S_t as $\pi|\tau|=\frac{2\pi}{3}|\lambda|$)

$$\int\limits_{S_t} |\mu(z)|\rho^2(z)\,dxdy \le C_6\big(k\delta+\frac{O_{\tau}(\mathbf{r})}{k}+\frac{N_{\tau}(\delta)}{|\lambda|}O_{\tau}(\mathbf{r})\big),$$

where C_6 is a constant that only depends on $||\mathbf{r}||_{\infty}$. Applying Lemma 2.1 and the last estimate we conclude that there is a constant C_7 that depends only on $||\mathbf{r}||_{\infty}$ such that

$$||\frac{\partial\psi}{\partial t}(t_0)||^2_{WP} \le C_7^2 O_{\tau}(\mathbf{r})\left(k\delta+\frac{O_{\tau}(\mathbf{r})}{k}+\frac{N_{\tau}(\delta)}{|\lambda|}O_{\tau}(\mathbf{r})\right).$$

Together with (18) this proves the estimate (16) and this theorem. □

The following theorem is the only result regarding Weil-Petersson distances that will be used later in this paper. We use the notation from the previous theorem.

Theorem 2.2. *Let $r_0>0$ and assume that for every $r>r_0$ we are given a finite type surface $S(r)$ (with at least one puncture) with an ideal triangulation $\tau(r)$ (the corresponding set of edges is $\lambda(r)$) that determines a point $\mathbf{r}(r)\in X(\tau(r))$ with the following properties*

- *There exists a universal constant $C > 0$ such that $\{||\mathbf{r}_m||_\infty\} \leq C$ for every $r > r_0$.*
- *Let $\delta(r) = re^{-r}$. Then there exists a polynomial $P(r)$ such that*
$$\frac{N_{\tau(r)}(\delta(r))}{|\lambda(r)|} \leq P(r)e^{-r}.$$
- *We have $O_{\tau(r)}(\mathbf{r}(r)) \leq 2r^2$.*

Then for some polynomial $P_1(r)$ we have
$$d_{WP}(F_{\tau(r)}(0), F_{\tau(r)}(\mathbf{r}(r))) \leq P_1(r)e^{-\frac{r}{4}}.$$
In particular the Weil-Petersson distance between the points $F_{\tau(r)}(0)$ and $F_{\tau(r)}(\mathbf{r}(r))$ tends to zero when $r \to \infty$.

Proof. Let k be an integer that is given by
$$e^{\frac{r}{2}} \leq k < e^{\frac{r}{2}} + 1.$$
Then $k \geq 2$ and $k\delta(r) < 1$ for r large enough. The proof follows directly from (16) for this choice of k.

□

3. Constructing finite degree covers of a punctured Riemann surface

3.1. **Admissible collections of triangles.** The aim of this section is to describe how to construct finite degree covers of a given punctured Riemann surface of finite type. We also need to keep the track of the geometry of these covers which is the motivation for the construction below. From now until the end of the paper S denotes a fixed Riemann surface of type $(\mathbf{g}, \mathbf{n})$ (we assume that $\mathbf{n} \geq 1$). The corresponding set of punctures in the boundary of S (or simply cusps) is denoted by $\text{Cusp}(S)$. We number the cusps in $\text{Cusp}(S)$ as $\text{Cusp}(S) = \{c_1(S), ..., c_{\mathbf{n}}(S)\}$. We also fix an ideal geodesic triangulation $\tau(S)$ on S where $\lambda(S)$ denotes the corresponding set of the edges. Most estimates we obtain will depend on the choice of S and $\tau(S)$.

Let $\Gamma(S)$ denote the set of all immersed ideal geodesics on S (all geodesics are taken with respect to the underlining hyperbolic metric). A geodesic γ on S is said to be ideal if both of its endpoints are in the set $\text{Cusp}(S)$ (note that these geodesics can have self intersections). We will say that a triangle T on S is an immersed ideal triangle if it lifts to an ideal triangle in the universal cover. The set of all immersed ideal triangles is denoted by $\mathcal{T}(S)$. If $T \in \mathcal{T}(S)$ then its edges belong to $\Gamma(S)$.

A geodesic $\gamma \in \Gamma(S)$ does not carry a natural orientation. That is, we have a choice of two orientations on each such γ. By $\Gamma^*(S)$ we denote the corresponding set of oriented geodesics. In particular, the set of oriented geodesics from $\lambda(S)$ will be called $\lambda^*(S)$. For $\gamma^* \in \Gamma^*(S)$ the corresponding unoriented geodesic is typically denoted by $\gamma \in \Gamma(S)$.

Let $\mathbf{N}\mathcal{T}(S)$ denote the space of all formal sums of triangles from $\mathcal{T}(S)$ over non-negative integers. Similarly, by $\mathbf{Z}\Gamma^*(S)$ we denote the space of all formal sums of oriented geodesics over all integers subject to the following rule. If γ_1^* and γ_2^* represent the same geodesic with opposite orientations then we have the identity $\gamma_1^* = -\gamma_2^*$.

We define the boundary operator $\partial : \mathbf{N}\mathcal{T}(S) \to \mathbf{Z}\Gamma^*(S)$ as follows. Let $T \in \mathcal{T}(S)$ and let γ_i, $i = 1, 2, 3$, denote its edges. Then $\partial T = \gamma_1^* + \gamma_2^* + \gamma_3^*$ where γ_i^* are the corresponding oriented geodesics so that the triangle T is to the left of each γ_i^*. In other words, the orientation of γ_i^* agrees with the orientation that γ_i inherits as the part of the boundary of T (such γ_i^* is called an oriented edge of the triangle T). The operator ∂ is extended to $\mathbf{N}\mathcal{T}$ by linearity.

Note that the space $\mathbf{N}\mathcal{T}(S)$ naturally embeds into the corresponding group of 2-chains on S. Also, $\mathbf{Z}\Gamma^*(S)$ naturally embeds into the corresponding group of 1-chains on S. The operator ∂ we defined agrees with the restriction of the standard boundary operator on $\mathbf{N}\mathcal{T}(S)$. Let $H_1(\overline{S}, \operatorname{Cusp}(S))$ denote the first homology of the surface $\overline{S}$ relative to the boundary of S which is the set of cusps $\operatorname{Cusp}(S)$. The class of each element of $\mathbf{Z}\Gamma^*(S)$ represents an element in $H_1(\overline{S}, \operatorname{Cusp}(S))$. In particular, for every $R \in \mathbf{N}\mathcal{T}(S)$ we have that ∂R is equal to zero in $H_1(\overline{S}, \operatorname{Cusp}(S))$ (although ∂R is not necessarily equal to zero in $\Gamma^*(S)$).

By G we always denote a Fuchsian group so that S is isomorphic to $\mathbf{H}/G$. The lift of $\tau(S)$, $\lambda(S)$, $\lambda^*(S)$, $\mathcal{T}(S)$, $\Gamma(S)$ and $\Gamma^*(S)$ under this uniformisation will be denoted by $\tau(G)$, $\lambda(G)$, $\lambda^*(G)$, $\mathcal{T}(G)$, $\Gamma(G)$ and $\Gamma^*(G)$. Also, by $\operatorname{Cusp}(G)$ we denote the set of cusps in the boundary of $\mathbf{H}$ (we identify $\partial\mathbf{H}$ with $\overline{\mathbf{R}}$), that is, $\operatorname{Cusp}(G)$ is the set of fixed points of all parabolic elements in G. We can identify the quotient $\operatorname{Cusp}(G)/G$ with the set of punctures in the boundary of S which we denoted by $\operatorname{Cusp}(S)$. If $T \in \mathcal{T}(G)$ by $[T]_G$ we denote the orbit of T under G. We identify $[T]_G$ with the corresponding element of $\mathcal{T}(S)$.

The set $\mathcal{T}^*(G)$ is defined as the set of pairs (T, γ^*) where $T \in \mathcal{T}(G)$ and $\gamma^* \in \Gamma^*(G)$ is an oriented edge of T. If $g \in G$ then $(g(T), g(\gamma^*)) \in \mathcal{T}^*(G)$ as well. Set $\mathcal{T}^*(S) = \mathcal{T}^*(G)/G$ (this definition does not depend on the choice of the group G). The equivalence class of (T, γ^*) is denoted by $[(T, \gamma^*)]_G \in \mathcal{T}^*(S)$. Set $\operatorname{Pr}_S([(T, \gamma^*)]_G) = [T]_G \in \mathcal{T}(S)$. Then $\operatorname{Pr}_S : \mathcal{T}^*(S) \to \mathcal{T}(S)$ is well defined.

Remark. Let $T \in \mathcal{T}(S)$ and let $\gamma_1^*, \gamma_2^*, \gamma_3^*$ denote its oriented edges. It can happen that $\gamma_1^* = \gamma_2^*$. This is why we can not define $\mathcal{T}^*(S)$ as the set of pairs (T, γ^*) where $T \in \mathcal{T}(S)$ and $\gamma^* \in \Gamma^*(S)$ is an oriented edge of T.

If $(T_1, \gamma^*), (T_2, -\gamma^*) \in \mathcal{T}^*(G)$ then $[(T_1, \gamma^*)]_G \neq [(T_2, -\gamma^*)]$. To prove this, assume that $[(T_1, \gamma^*)]_G = [(T_2, -\gamma^*)]$. Then there is $g \in G$ so that $g(\gamma^*) = -\gamma^*$. Also, such g is not the identity map. But then $g^2(\gamma^*) = \gamma^*$. This implies that g fixes the endpoints $c_1, c_2 \in \operatorname{Cusp}(G)$ of γ^*. Since $g^2(c_1) = c_1$ we have that g is a parabolic element of G and therefore c_1 is the unique fixed point of g^2. This contradicts the fact that $g^2(c_2) = c_2$.

Fix $T \in \mathcal{T}(G)$ and let $\mathrm{R}_T : \mathbf{H} \to \mathbf{H}$ be the standard rotation around the centre of T for the angle $2\pi/3$ (this rotation is of the order three). For $(T, \gamma^*) \in \mathcal{T}^*(G)$ set $\operatorname{rot}_G(T, \gamma^*) = (T, \mathrm{R}_T(\gamma^*)) \in \mathcal{T}^*(G)$. The map $\operatorname{rot}_G : \mathcal{T}^*(G) \to \mathcal{T}^*(G)$ is a bijection and rot_G^3 is the identity map. It is obvious that rot_G and rot_G^2 have no fixed points. If $g \in G$ then $g \circ \mathrm{R}_T = \mathrm{R}_{g(T)} \circ g$. This shows that the map $\operatorname{rot}_S : \mathcal{T}^*(S) \to \mathcal{T}^*(S)$ given by $\operatorname{rot}_S([(T, \gamma^*)]_G) = [(T, \mathrm{R}_T(\gamma^*))]_G$ is well defined. Assume that $\operatorname{rot}_S([(T, \gamma^*)]_G) = [(T, \gamma^*)]_G$. Then there is $g \in G$ so that $g(T) = T$ and so that $g(\gamma^*) = \mathrm{R}_T(\gamma^*)$. This implies that g and R_T agree at the vertices of T which shows that $g = \mathrm{R}_T$. This is not possible since G is torsion free. The conclusion is that the map rot_S has no fixed points. Since rot_S^3 is the identity map, we conclude that rot_S^2 has no fixed points either. Note that $\operatorname{Pr}_S \circ \operatorname{rot}_S = \operatorname{Pr}_S$.

Let $\mathcal{L}_{\mathcal{C}}$ be a finite set of labels, and let $\text{lab}_{\mathcal{C}} : \mathcal{L}_{\mathcal{C}} \to \mathcal{T}^*(S)$ be a labelling map. We say that the pair $\mathcal{C} = (\mathcal{L}_{\mathcal{C}}, \text{lab}_{\mathcal{C}})$ is a labelled collection of triangles if the following holds:

- There exists a bijection $\text{rot}_{\mathcal{C}} : \mathcal{L}_{\mathcal{C}} \to \mathcal{L}_{\mathcal{C}}$ so that $\text{rot}^3_{\mathcal{C}}$ is the identity map.
- We have $\text{lab}_{\mathcal{C}} \circ \text{rot}_{\mathcal{C}} = \text{rot}_S \circ \text{lab}_{\mathcal{C}}$.

Since rot_S and rot^2_S have no fixed points, we see that $\text{rot}_{\mathcal{C}}$ and $\text{rot}^2_{\mathcal{C}}$ have no fixed points either.

Choose $\gamma^* \in \Gamma^*(S)$ and let γ^*_1 be its lift to $\mathbf{H}$. We say that $a \in \mathcal{L}_{\mathcal{C},\gamma^*} \subset \mathcal{L}_{\mathcal{C}}$ if $\text{lab}_{\mathcal{C}}(a) = [(T, \gamma^*_1)]_G$ for some $T \in \mathcal{T}(G)$ so that $(T, \gamma^*_1) \in \mathcal{T}^*(G)$. Then $\sigma_{\mathcal{C}}(\mathcal{L}_{\mathcal{C},\gamma^*}) = \mathcal{L}_{\mathcal{C},-\gamma^*}$. It follows from the above discussion that the sets $\mathcal{L}_{\mathcal{C},\gamma^*}$ and $\mathcal{L}_{\mathcal{C},-\gamma^*}$ are disjoint (if the were not there would be an order two element in G as we showed above). Let $\gamma_i \in \Gamma(S)$, $i = 1, ..., k$, be the set of different edges of triangles from $\text{Pr}_S(\text{lab}_{\mathcal{C}})(\mathcal{L}_{\mathcal{C}})$. Choose an orientation for γ^*_i for each γ_i. Then $\mathcal{L}_{\mathcal{C}}$ is the disjoint union $\mathcal{L}_{\mathcal{C}} = \mathcal{L}_{\mathcal{C},\gamma^*_1} \cup \mathcal{L}_{\mathcal{C},-\gamma^*_1} \cup ... \cup \mathcal{L}_{\mathcal{C},\gamma^*_k} \cup \mathcal{L}_{\mathcal{C},-\gamma^*_k}$.

Set $(\text{Pr}_S \circ \text{lab}_{\mathcal{C}})(\mathcal{L}_{\mathcal{C}}) = \{T_1, T_2, ..., T_m\} \subset \mathcal{T}(S)$. Let $k_i = |(\text{Pr}_S \circ \text{lab}_{\mathcal{C}})^{-1}(T_i)|$, where $|(\text{Pr}_S \circ \text{lab}_{\mathcal{C}})^{-1}(T_i)|$ is the number of elements in the preimage $(\text{Pr}_S \circ \text{lab}_{\mathcal{C}})^{-1}(T_i)$. Since $\text{rot}_{\mathcal{C}}$ has no fixed points, and since $\text{Pr}_S \circ \text{rot}_S = \text{Pr}_S$ we have that $k_i = 3l_i$, for some integer l_i. Then each $\mathcal{C}$ induces an element in $\mathbf{N}\mathcal{T}(S)$ given by $k_1T_1 + kT_2 + ... + k_mT_m = 3l_1T_1 + 3l_2T_2 + ... + 3l_mT_m$. If there is no confusion we will denote this element of $\mathbf{N}\mathcal{T}(S)$ by $\mathcal{C}$ as well. We define $\partial\mathcal{C}$ as the boundary of the corresponding element of $\mathbf{N}\mathcal{T}(S)$.

Remark. Every $R \in \mathbf{N}\mathcal{T}(S)$ where $R = l_1T_1 + l_2T_2 + ... + l_mT_m$ and l_i are positive integers, induces a labelled collection of triangles as follows (this will be used in the proof of Theorem 1.1 in the next section). The corresponding set of labels is $\mathcal{L}_{\mathcal{C}} = \{(i, i') : i = 1, 2, ..., (l_1 + l_2 + ... + l_m); i' = 1, 2, 3\}$ (note that the set $\mathcal{L}_{\mathcal{C}}$ has $3(l_1 + ... + l_m)$ elements). Let $T'_j \in \mathcal{T}(G)$ be a lift of T_j. Let $\gamma^*_{i'}$, $i' = 1, 2, 3$, be its oriented edges, so that $\text{R}_T(\gamma^*_1) = \gamma^*_2$. The corresponding labelling map $\text{lab}_{\mathcal{C}}$ is given by $\text{lab}_{\mathcal{C}}(i, i') = [(T'_j, \gamma^*_{i'})]_G$, for $l_1 + ... + l_{j-1} < i \le l_1 + ... + l_j$. The required bijection $\text{rot}_{\mathcal{C}} : \mathcal{L}_{\mathcal{C}} \to \mathcal{L}_{\mathcal{C}}$ is given by $\text{rot}_{\mathcal{C}}(i, j) = (i, j + 1 \mod 3)$.

Definition 3.1. *Let $\mathcal{C}$ be a labelled collection of triangles. Let $\sigma_{\mathcal{C}} : \mathcal{L}_{\mathcal{C}} \to \mathcal{L}_{\mathcal{C}}$ be a bijection. We say that the pair $(\mathcal{C}, \sigma_{\mathcal{C}})$ is an admissible pair if the following holds:*

- *The map $\sigma_{\mathcal{C}}$ is an involution, that is, $\sigma^2_{\mathcal{C}}$ is the identity map, and $\sigma_{\mathcal{C}}$ has no fixed points.*
- *If $a, b \in \mathcal{L}_{\mathcal{C}}$ and $\sigma_{\mathcal{C}}(a) = b$ then there are triangles $T_1, T_2 \in \mathcal{T}(G)$ and $\gamma^* \in \Gamma^*(G)$ such that $\text{lab}_{\mathcal{C}}(a) = [(T_1, \gamma^*)]_G$ and $\text{lab}_{\mathcal{C}}(b) = [(T_2, -\gamma^*)]_G$ (recall that $-\gamma^*$ denotes the opposite orientation of γ^*).*

If $(\mathcal{C}, \sigma_{\mathcal{C}})$ is an admissible pair, we have two bijections $\sigma_{\mathcal{C}}, \text{rot}_{\mathcal{C}} : \mathcal{L}_{\mathcal{C}} \to \mathcal{L}_{\mathcal{C}}$. The group generated by these two bijections is denoted by $\langle \sigma_{\mathcal{C}}, \text{rot}_{\mathcal{C}} \rangle$.

One can construct examples of admissible pairs as follows. Let $S_1 \to S$ be a cover, where S_1 is a finite type surface (recall that in this paper all the covers, except the universal cover, are assumed to be finite degree, regular and holomorphic). Then there is a Fuchsian group $G_1 < G$ so that S_1 is isomorphic to $\mathbf{H}/G_1$. Let $\tau_1(S_1)$ be a geodesic triangulation of S_1 and let $\tau_1(G_1)$ denote its lift to $\mathbf{H}$. By $\tau^*_1(G_1)$ we denote the set of pairs (T, γ^*) where $T \in \tau_1(G_1)$ and γ^* is an oriented edge of T. Same as above, the group G_1 acts on $\tau^*_1(G_1)$. By $[(T, \gamma^*)]_{G_1}$ we denote the corresponding orbit, and by $\tau^*_1(S_1) = \tau^*_1(G_1)/G_1$ we denote the corresponding set of orbits. Same as above, we define the map $\text{rot}_{S_1} : \tau_1^*(S_1) \to \tau_1^*(S_1)$. Then rot_{S_1} is a bijection of order three.

Set $\mathcal{L}_{\mathcal{C}} = \tau_1^*(S_1)$ and let $\text{rot}_{\mathcal{C}} = \text{rot}_{S_1}$. For $[(T,\gamma^*)]_{G_1} \in \tau_1^*(S_1)$ set $\text{lab}_{\mathcal{C}}([(T,\gamma^*)]_{G_1}) = [(T,\gamma^*)]_G \in \mathcal{T}^*(S)$. One can verify that $\text{lab}_{\mathcal{C}} \circ \text{rot}_{\mathcal{C}} = \text{rot}_S \circ \text{lab}_{\mathcal{C}}$. Then $\mathcal{C} = (\mathcal{L}_{\mathcal{C}}, \text{lab}_{\mathcal{C}})$ is a labelled collection of triangles. We now define the involution $\sigma_{\mathcal{C}}$.

Let $a \in \mathcal{L}_{\mathcal{C}}$. We want to define $\sigma_{\mathcal{C}}(a) \in \mathcal{L}_{\mathcal{C}}$. Let $T \in \tau_1(G_1)$ and $\gamma^* \in \Gamma^*(G)$ be its oriented edge, so that $a = [(T,\gamma^*)]_{G_1}$. Let T_1 be the unique triangle in $\tau_1(G_1)$ that is adjacent to T along γ^*. Set $\sigma_{\mathcal{C}}(a) = [(T_1,-\gamma^*)]_{G_1}$. Clearly $(T_1,-\gamma^*) \in \tau_1^*(G_1)$ so $\sigma_{\mathcal{C}}(a)$ is well defined. Moreover, $\sigma_{\mathcal{C}}$ is of order two. If $\sigma_{\mathcal{C}}$ had a fixed point, then there would exist an element of order two in G which is not possible. We have that $(\mathcal{C},\sigma_{\mathcal{C}})$ is an admissible pair.

Definition 3.2. *The above constructed pair $(\mathcal{C},\sigma_{\mathcal{C}})$ is called a virtual triangulation pair.*

If $(\mathcal{C}(i),\sigma_{\mathcal{C}(i)})$, $i=1,...,m$, are virtual triangulations pairs, then we can construct a new admissible pair $(\mathcal{C},\sigma_{\mathcal{C}})$ as follows. Set $\mathcal{L}_{\mathcal{C}} = \mathcal{L}_{\mathcal{C}(1)} \cup ... \cup \mathcal{L}_{\mathcal{C}(m)}$ (here we assume that the sets of labels $\mathcal{L}_{\mathcal{C}(i)}$ are mutually disjoint). On each $\mathcal{C}(i)$ the map $\text{rot}_{\mathcal{C}}$ agrees with $\text{rot}_{\mathcal{C}(i)}$. Also, on each $\mathcal{L}_{\mathcal{C}(i)}$ the map $\text{lab}_{\mathcal{C}}$ agrees with the map $\text{lab}_{\mathcal{C}(i)}$. Then $\mathcal{C}$ is a labelled collection of triangles. We define $\sigma_{\mathcal{C}}$ to agree with $\sigma_{\mathcal{C}(i)}$ on each $\mathcal{L}_{\mathcal{C}(i)}$. We have that $(\mathcal{C},\sigma_{\mathcal{C}})$ is an admissible pair.

Definition 3.3. *Let $(\mathcal{C}(i),\sigma_{\mathcal{C}(i)})$, $i=1,...,m$, be admissible pairs. We say that an admissible pair $(\mathcal{C},\sigma_{\mathcal{C}})$ is the union of $(\mathcal{C}(i),\sigma_{\mathcal{C}(i)})$, $i=1,...,m$, if the following holds.*

- *There exist injections $\phi_i : \mathcal{L}_{\mathcal{C}(i)} \to \mathcal{L}_{\mathcal{C}}$ so that the set $\mathcal{L}_{\mathcal{C}}$ is the disjoint union of the sets $\phi_i(\mathcal{L}_{\mathcal{C}(i)})$.*
- *For each ϕ_i we have $\text{lab}_{\mathcal{C}} \circ \phi_i = \text{lab}_{\mathcal{C}(i)}$.*
- *We have $\phi \circ \text{rot}_{\mathcal{C}(i)} = \text{rot}_{\mathcal{C}} \circ \phi$ and $\phi \circ \sigma_{\mathcal{C}(i)} = \sigma_{\mathcal{C}} \circ \phi$.*

We have the following lemma.

Lemma 3.1. *For every admissible triangulation pair $(\mathcal{C},\sigma_{\mathcal{C}})$ there exist virtual triangulation pairs $(\mathcal{C}(i),\sigma_{\mathcal{C}(i)})$, $i=1,...,l$, so that $(\mathcal{C},\sigma_{\mathcal{C}})$ is the union of $(\mathcal{C}(i),\sigma_{\mathcal{C}(i)})$.*

Proof. We divide $\mathcal{L}_{\mathcal{C}}$ into the orbits of the group $\langle \sigma_{\mathcal{C}}, \text{rot}_{\mathcal{C}} \rangle$. Denote these orbits by $\mathcal{L}_{\mathcal{C}(i)}$ where i goes through some finite set of labels. Let $\text{lab}_{\mathcal{C}(i)}$, $\sigma_{\mathcal{C}(i)}$, and $\text{rot}_{\mathcal{C}(i)}$ denote the restrictions of the the maps $\text{lab}_{\mathcal{C}}$, $\sigma_{\mathcal{C}}$, and $\text{rot}_{\mathcal{C}}$ on the set $\mathcal{L}_{\mathcal{C}(i)}$. Then each pair $\mathcal{C}(i) = (\mathcal{L}_{\mathcal{C}(i)}, \text{lab}_{\mathcal{C}(i)})$ is an admissible pair, and $(\mathcal{C},\sigma_{\mathcal{C}})$ is the union of $(\mathcal{C}(i),\sigma_{\mathcal{C}(i)})$. Moreover, the group $\left\langle \sigma_{\mathcal{C}(i)}, \text{rot}_{\mathcal{C}(i)} \right\rangle$ acts transitively on $\mathcal{L}_{\mathcal{C}(i)}$. Therefore, in order to prove the lemma, it suffices to prove that every admissible pair $(\mathcal{C},\sigma_{\mathcal{C}})$ for which the group $\langle \sigma_{\mathcal{C}}, \text{rot}_{\mathcal{C}} \rangle$ acts transitively on $\mathcal{L}_{\mathcal{C}}$ is in fact a virtual triangulation pair.

Assume that $(\mathcal{C},\sigma_{\mathcal{C}})$ is an admissible pair, so that the group $\langle \sigma_{\mathcal{C}}, \text{rot}_{\mathcal{C}} \rangle$ acts transitively on $\mathcal{L}_{\mathcal{C}}$. Let $\mathcal{T}^*_{\mathcal{C}}(G)$ be the space of triples (T,γ^*,a) where $(T,\gamma^*) \in \mathcal{T}^*(G)$ and where $\text{lab}_{\mathcal{C}}(a) = [(T,\gamma^*)]$. There is a unique involution $\sigma_{\mathcal{C}}(G) : \mathcal{T}^*_{\mathcal{C}}(G) \to \mathcal{T}^*_{\mathcal{C}}(G)$ such that $\sigma_{\mathcal{C}}(G)((T_1,\gamma_1^*,a_1)) = (T_2,\gamma_2^*,a_2)$ if and only if $-\gamma_1^* = \gamma_2^*$ and $a_1 = \sigma_{\mathcal{C}}(a_2)$. Define $\text{rot}_{\mathcal{C}}(G) : \mathcal{T}^*_{\mathcal{C}}(G) \to \mathcal{T}^*_{\mathcal{C}}(G)$ by $\text{rot}_{\mathcal{C}}(G)((T,\gamma^*,a)) = (T, \text{R}_T(\gamma^*), \text{rot}_{\mathcal{C}}(a))$. It is easy to verify that the group $\langle \sigma_{\mathcal{C}}(G), \text{rot}_{\mathcal{C}}(G) \rangle$ is isomorphic to $\mathbf{Z}_2 \star \mathbf{Z}_3$ and that it acts freely on $\mathcal{T}^*_{\mathcal{C}}(G)$. Moreover, the collection of triangles from $\mathcal{T}(G)$ that appears in a given orbit under this action, is an ideal triangulation of $\mathbf{H}$.

The group G naturally acts on $\mathcal{T}^*_{\mathcal{C}}(G)$ by $g(T,\gamma^*,a)=(g(T),g(\gamma^*),a)\in\mathcal{T}^*_{\mathcal{C}}(G)$ (note that $[(T,\gamma^*)]=[(g(T),g(\gamma^*))]$). If $(T_1,\gamma^*_1,a),(T_2,\gamma^*_2,a)\in\mathcal{T}^*_{\mathcal{C}}(G)$ then by definition we have $[(T_1,\gamma^*_1)]=[(T_2,\gamma^*_2)]$. That is, there is $g\in G$ so that $g(T_1,\gamma^*_1,a)=(T_1,\gamma^*_1,a)$. In particular, this shows that the group generated by G and $\langle\sigma_{\mathcal{C}}(G),\mathrm{rot}_{\mathcal{C}}(G)\rangle$ acts transitively on $\mathcal{T}^*_{\mathcal{C}}(G)$. Also, every element from G commutes with every element from $\langle\sigma_{\mathcal{C}}(G),\mathrm{rot}_{\mathcal{C}}(G)\rangle$.

Fix $(T,\gamma^*,a)\in\mathcal{T}^*_{\mathcal{C}}(G)$ and consider the orbit $O=\langle\sigma_{\mathcal{C}}(G),\mathrm{rot}_{\mathcal{C}}(G)\rangle\,(T,\gamma^*,a)$. Let G_1 be the subgroup of G so that $g\in G_1$ if g preserves the orbit O. If $(T_1,\gamma^*_1,a),(T_2,\gamma^*_2,a)\in O$ then there is $g\in G$ so that $g(T_1,\gamma^*_1,a)=(T_1,\gamma^*_1,a)$. Moreover, since g commutes with elements from $\langle\sigma_{\mathcal{C}}(G),\mathrm{rot}_{\mathcal{C}}(G)\rangle$ we conclude that g preserves the entire orbit O. This shows that the map $\psi:O/G_1\to\mathcal{L}_{\mathcal{C}}$ given by $\phi(T,\gamma^*,a)=a$ is an injection. Since the group $\langle\sigma_{\mathcal{C}},\mathrm{rot}_{\mathcal{C}}\rangle$ acts transitively on $\mathcal{L}_{\mathcal{C}}$ we find that ψ is a bijection.

Let S_1 be the Riemann surface isomorphic to $\mathbf{H}/G_1$. Set $\tau^*(S_1)=O/G_1$ and denote by $\tau(S_1)$ the union of triangles that appear in O/G_1. Then $\tau(S_1)$ is an ideal triangulation of the surface S_1. Since $\tau(S_1)$ contains finitely many such triangles (the map $\psi:\tau^*(S_1)\to\mathcal{L}_{\mathcal{C}}$ is a bijection onto a finite set $\mathcal{L}_{\mathcal{C}}$), we see that S_1 is a finite type surface, and the group G_1 has finite index in G. Now it follows that the admissible pair $(\mathcal{C},\sigma_{\mathcal{C}})$ is a virtual triangulation pair, that via the map ψ corresponds to the triangulation of S_1.

□

Consider a labelled collection of triangles $\mathcal{C}$ as an element of $\mathbf{N}\mathcal{T}(S)$. Assume that $(\mathcal{C},\sigma_{\mathcal{C}})$ is an admissible pair. We saw above that $\mathcal{L}_{\mathcal{C}}$ is the disjoint union $\mathcal{L}_{\mathcal{C}}=\mathcal{L}_{\mathcal{C},\gamma^*_1}\cup\mathcal{L}_{\mathcal{C},-\gamma^*_1}\cup...\cup\mathcal{L}_{\mathcal{C},\gamma^*_k}\cup\mathcal{L}_{\mathcal{C},-\gamma^*_k}$ where $\gamma_i\in\Gamma^*(S)$ are the different edges of triangles from $\mathrm{Pr}_S(\mathrm{lab}_{\mathcal{C}}(\mathcal{L}_{\mathcal{C}}))$. Then $\sigma_{\mathcal{C}}(\mathcal{L}_{\mathcal{C},\gamma^*_i})=\mathcal{L}_{\mathcal{C},-\gamma^*_i}$. We conclude that the sets $\mathcal{L}_{\mathcal{C},\gamma^*_i}$ and $\mathcal{L}_{\mathcal{C},-\gamma^*_i}$ have the same number of elements, which implies that γ^*_i does not figure in $\partial\mathcal{C}$. We have that $\partial\mathcal{C}$ is equal to zero in the space $\mathbf{Z}\Gamma^*(S)$.

On the other hand, if $\mathcal{C}$ is a labelled collection of triangles so that $\partial\mathcal{C}$ is zero in $\mathbf{Z}\Gamma^*(S)$ then the corresponding sets $\mathcal{L}_{\mathcal{C},\gamma^*}$ and $\mathcal{L}_{\mathcal{C},-\gamma^*}$ have the same number of elements. Again, from the fact that $\mathcal{L}_{\mathcal{C}}$ is the disjoint union $\mathcal{L}_{\mathcal{C}}=\mathcal{L}_{\mathcal{C},\gamma^*_1}\cup\mathcal{L}_{\mathcal{C},-\gamma^*_1}\cup...\cup\mathcal{L}_{\mathcal{C},\gamma^*_k}\cup\mathcal{L}_{\mathcal{C},-\gamma^*_k}$ we can construct an appropriate involution $\sigma_{\mathcal{C}}$. However, there are many ways in which we can do this (unless each $\mathcal{L}_{\mathcal{C},\gamma^*_i}$ has one element). This is where the geometric considerations start. Our aim will be to construct these involutions so that the pair of triangles that are related by the involution satisfies that the hyperbolic distance between the orthogonal projections of the centres of these triangles to their common edge, is as small as possible.

3.2. **The definitions of height, 0-horoball and combinatorial length.** As we said, G denotes a Fuchsian group so that S is isomorphic to $\mathbf{H}/G$. For every $c\in\mathrm{Cusp}(G)$ by $G(c)$ we denote the orbit of c under the group G. Each such orbit corresponds to a cusp $c_i(S)\in\mathrm{Cusp}(S)$. Fix $c_i(S)$ that is we fix the orbit of some $c\in\mathrm{Cusp}(G)$. Let $f:\mathbf{H}\to\mathbf{H}$ be a Möbius transformation, so that $f(c)=\infty$ and so that the parabolic element in in $f\circ G\circ f^{-1}$ that generates the corresponding cyclic subgroup of $f\circ G\circ f^{-1}$, is the translation $g_\infty(z)=z+1$ (there are many such Möbius transformations f but we choose one for each orbit). The group $G_{c_i}=G_c=f\circ G\circ f^{-1}$ is called the normalised group with respect to the cusp c. If $c'\in G(c)$ then $G_{c_i(S)}=G_c=G_{c'}$. There are exactly $\mathbf{n}$ normalised groups

G_c (where $\mathbf{n}$ is the number of punctures in the boundary of S). From now on G is one of the normalised groups (if we do not specify which one, then it can be any one). The constants we introduce below may depend on the choice of the group G. However, since we consider only finitely many such groups, these constants depend only on S.

Definition 3.4. *Let $z \in \mathbf{H}$ $c \in \operatorname{Cusp}(G)$. Let f be the corresponding Möbius transformation that conjugates G onto G_c. We have:*

- *Set $\mathbf{h}_c(z) = \log(\operatorname{Im}(f(z)))$ where $\operatorname{Im}(f(z))$ is the imaginary part of $f(z) \in \mathbf{H}$. We say that $\mathbf{h}_c(z)$ is the height of the point z with respect to the cusp $c \in \operatorname{Cusp}(G)$.*
- *Let $T \in \mathcal{T}(G)$ and denote by $c_1, c_2, c_3 \in \operatorname{Cusp}(G)$ its vertices. Set $\mathbf{h}(T) = \max\{|\mathbf{h}_{c_1}(\operatorname{ct}(T))|, |\mathbf{h}_{c_2}(\operatorname{ct}(T))|, |\mathbf{h}_{c_3}(\operatorname{ct}(T))|\}$. Note that $h(T) \geq 0$.*
- *Set $\mathbf{h}(z) = \max_{c \in \operatorname{Cusp}(G)} \mathbf{h}_c(z)$.*
- *Let $c \in \operatorname{Cusp}(G)$ and let $\gamma \in \Gamma(G)$ be such so that c is not one of its endpoints. Denote by $\mathbf{z}_{\max}(\gamma, c)$ the point on γ so that $\mathbf{h}_c(\mathbf{z}_{\max}(\gamma, c)) \geq \mathbf{h}_c(z)$ for any other point $z \in \gamma$.*
- *Let $t \in \mathbf{R}$. Set $\mathcal{H}_c(t) = \{z \in \mathbf{H} : \mathbf{h}_c(z) \geq t\}$. We say that $\mathcal{H}_c(t)$ is the t-horoball at c.*
- *Let $t \in \mathbf{R}$. We set $\mathbf{Th}_G(t) = \{z \in \mathbf{H} : \mathbf{h}(z) \leq t\}$ and $\mathbf{Thin}_G(t) = \mathbf{H} \setminus \mathbf{Th}_G(t)$.*

Remark. We establish the following related definitions. If T_1, T_2 are two triangles so that $f(T_1) = T_2$ for some $f \in G$ then $\mathbf{h}(T_1) = T_2$. This shows that $\mathbf{h}(T)$ is well defined for every $T \in \mathcal{T}(S)$. If c_1, c_2 are equivalent under G then the projection of the horoballs $\mathcal{H}_{c_1}(t)$ and $\mathcal{H}_{c_2}(t)$ to S agree. Therefore we can define the t-horoball $\mathcal{H}_c(t) \subset S$ for every $c \in \operatorname{Cusp}(S)$. It is well known that any two 0-horoballs on S either coincide or are disjoint (note that the closures of two zero horoballs for $\mathbf{PSL}(2, \mathbf{Z})$ may touch, but if a Fuchsian group G is a covering group of a Riemann surface, then this can not happen) . Therefore, for any $z \in \mathbf{H}$ there can exist at most one cusp $c \in \operatorname{Cusp}(G)$ so that $\mathbf{h}_c(z) > 0$. Moreover, given $z \in \mathbf{H}$ and any interval $[a, \infty) \subset \mathbf{R}$ there are at most finitely many cusps $c \in \operatorname{Cusp}(G)$ so that $\mathbf{h}_c(z) \in [a, b]$. This shows that $\mathbf{h}(z)$ is well defined. If $g \in G$ then $\mathbf{h}(z) = \mathbf{h}(g(z))$. This shows that for $p \in S$ the value $\mathbf{h}(p)$ is well defined. Also, by $\mathbf{Th}_S(t) = \{p \in S : \mathbf{h}(p) \leq t\}$ we denote the projection of the set $\mathbf{Th}_G(t) = \{z \in \mathbf{H} : \mathbf{h}(z) \leq t\}$ to S. We have $\mathbf{Thin}_S(t) = S \setminus \mathbf{Th}_S(t)$.

In the following proposition we introduce certain constants that will be used in the proof of the correction lemma below. The constant $N(S)$ will be used throughout the paper. Since S is of finite type the following numbers are well defined.

Proposition 3.1. *There exist a large enough number $t_1(S) > 0$ a sufficiently negative number $t_2(S) < 0$ and an integer $N(S) \geq 1$ with the following properties:*

- *Let $c \in \operatorname{Cusp}(G)$ and let $N_c(S)$ denote the number of non-equivalent geodesics from $\lambda(G)$ that end at c. Then the integer $N(S) = \max_{c \in \Gamma(G)} N_c(S)$ is well defined.*
- *Let $c \in \operatorname{Cusp}(G)$ and let $z \in \mathcal{H}_c(-\log(8N(S)))$. Then $t_1(S)$ is large enough so that z does not belong to any $t_1(S)$-horoball, except possibly the one at c.*
- *We have $e^{t_1(S)} > 4N(S)$ and $e^{t_1(S)} > 4 + 2N(S) + 1$.*

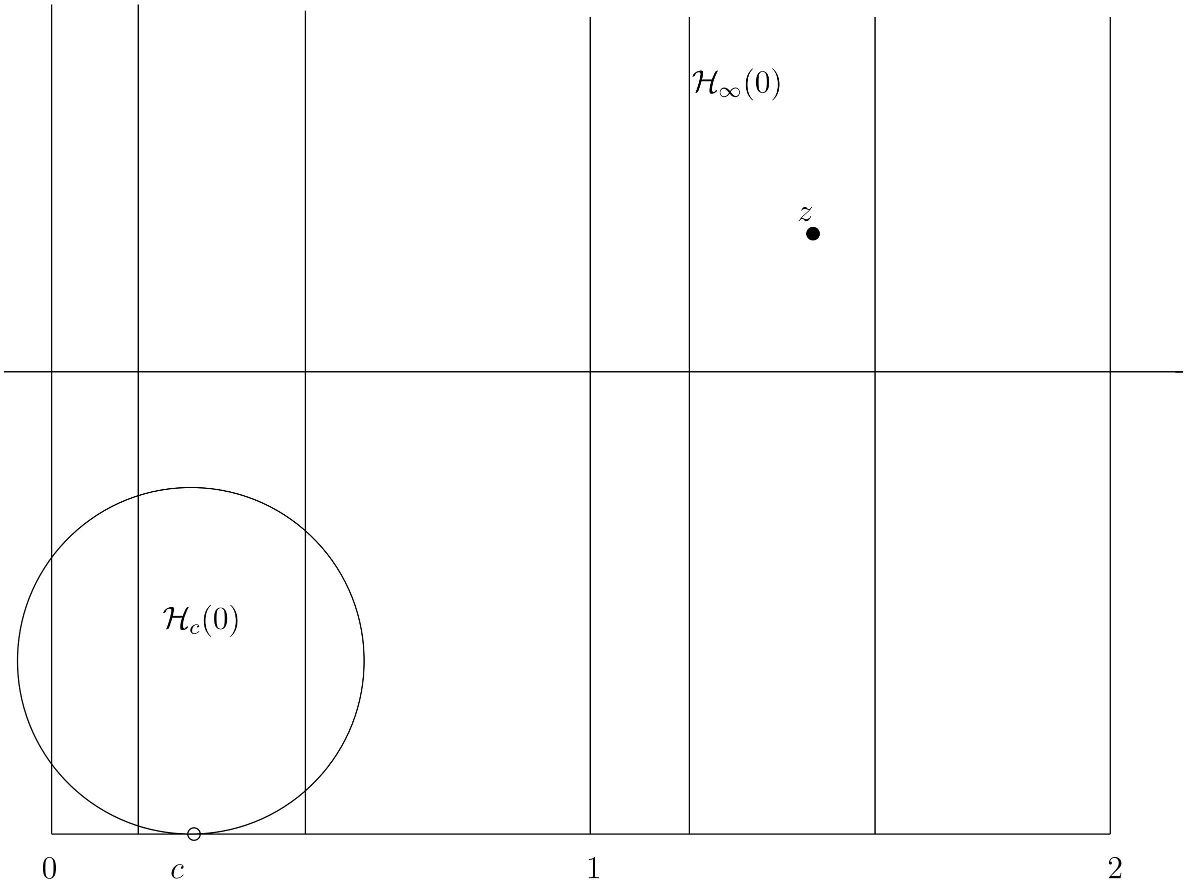

FIGURE 2. The case $N(S) = 3$; $h_\infty(z) = h(z) > 0$

- *Let $\gamma \in \lambda(G)$ and let $c_1, c_2 \in \mathrm{Cusp}(G)$ be its endpoints. Let $T \in \tau(G)$ be one of the two triangles that have γ in their boundary. Then the number $t_2(S)$ is sufficiently negative, so that for every point $z \in \mathbf{Th}_G(t_1(S))$ and $z \in \gamma$ we have $z \in \mathcal{H}_{c_1}(t_2(S))$ and $z \in \mathcal{H}_{c_2}(t_2(S))$.*

Remark. If $t > t'$ then for every $c \in \mathrm{Cusp}(G)$ we have $\mathcal{H}_c(t) \subset \mathcal{H}_c(t')$.

Proof. There are only $|\lambda(S)|$ non-equivalent geodesics in $\lambda(G)$. Also, the set $\gamma \cap \mathbf{Th}_G$ is compact in $\mathbf{H}$. The propositions follows from the basic compactness argument. □

If $\gamma \subset \mathbf{H}$ is an arbitrary geodesic segment (or finite or infinite length), by $\iota(\gamma, \tau(G))$ we denote the number of (transverse) intersections between γ and edges from $\lambda(G)$. In particular, if $\gamma \in \Gamma(G)$ the intersection number $\iota(\gamma, \tau(G))$ is finite. For a geodesic segment $\gamma \subset S$ by $\iota(\gamma, \tau(S))$ we denote the number of (transverse) intersections between γ and edges from $\lambda(S)$.

Definition 3.5. *Let* $\gamma \in \Gamma(G)$ *and let* $z \in \gamma$*. Let* $c_1, c_2 \in \operatorname{Cusp}(G)$ *be the endpoints of* γ*. We define the combinatorial length* $\mathcal{K}(\gamma, z)$ *as*

$$\mathcal{K}(\gamma, z) = \iota(\gamma, \tau(G)) + \psi(c_1, c_2, z),$$

where $\psi(c_1, c_2, z) = 0$ *if* $\max\{\mathbf{h}_{c_1(z)}, \mathbf{h}_{c_2(z)}\} \le t_1(S)$ *and*

$$\psi(c_1, c_2, z) = \max\{[e^{\mathbf{h}_{c_1(z)}}], [e^{\mathbf{h}_{c_2(z)}}]\},$$

if $\max\{\mathbf{h}_{c_1(z)}, \mathbf{h}_{c_2(z)}\} > t_1(S)$*. Here* $[e^{\mathbf{h}_{c_1(z)}}]$ *and* $[e^{\mathbf{h}_{c_2(z)}}]$ *denote the integer parts of* $e^{\mathbf{h}_{c_1(z)}}$ *and* $e^{\mathbf{h}_{c_2(z)}}$ *respectively. Note that if* $z \in \mathbf{Th}_G(t_1(S))$ *then* $\mathcal{K}(\gamma, z) = \iota(\gamma, \tau(G))$*.*

Since the 0-horoballs are disjoint, only one of the numbers $\mathbf{h}_{c_1(z)}, \mathbf{h}_{c_2(z)}$ can be non-negative for any $z \in \mathbf{H}$. If $\gamma^* \in \Gamma^*(G)$ represents γ together with a choice of orientation on γ then we set $\mathcal{K}(\gamma^*, z) = \mathcal{K}(\gamma, z)$.

The following lemma will be used in the proof of Theorem 1.1 (in the next section).

Lemma 3.2. *Let* R *be a finite type Riemann surface that covers* S *and let* $\tau(R)$ *be a geodesic triangulation on* R *(by* $\lambda(R)$ *we denote the corresponding set of edges). Let* $\mathbf{r} \in X(\tau(R))$ *be the vector so that* R *is the underlying Riemann surface for the point* $F_{\tau(R)}(\mathbf{r})$ *(see Section 2.). Let* $A = \max_{T \in \tau(R)} |\mathbf{h}(T)|$*. Then* $O_{\tau(R)}(\mathbf{r}) \le 2A$*.*

Proof. Let G_1 be a finite index subgroup of G so that R is isomorphic to $\mathbf{H}/G_1$. By $\tau(G_1)$ and $\lambda(G_1)$ we denote the corresponding lifts of $\tau(R)$ and $\lambda(R)$. We have $\tau(G_1) \subset \mathcal{T}(G)$ so we can define the height $\mathbf{h}(T)$ for every $T \in \tau(G_1)$ (and by projecting T to R we define $\mathbf{h}(T)$ for $T \in \tau(R)$).

The sets of cusps for S and R agree. Choose $c \in \operatorname{Cusp}(G)$ and let G be normalised, and $c = \infty$. Let $\lambda_1, ..., \lambda_k \in \lambda(G_1)$ be a k-tuple of consecutive edges that all end at ∞. Let $\mathbf{r}_i \in \mathbf{R}$ denote the corresponding shear coordinates on λ_i $i = 1, ..., k$. Let $T_i \in \tau(G_1)$ be the triangle that has λ_i as its edge, and so that T_i is to the left of λ_i^* where the orientation λ_i^* is chosen so that $+\infty$ is to the right of λ_i^*. Then

$$\operatorname{Im}(\operatorname{ct}(T_k)) = e^{\mathbf{r}_1 + \cdots + \mathbf{r}_{k-1}} \operatorname{Im}(\operatorname{ct}(T_1)),$$

and

$$|\mathbf{r}_1 + ... + \mathbf{r}_{k-1}| = |\log(\operatorname{Im}(\operatorname{ct}(T_k))) - \log(\operatorname{Im}(\operatorname{ct}(T_1)))| \le |\log(\operatorname{Im}(\operatorname{ct}(T_k)))| + |\log(\operatorname{Im}(\operatorname{ct}(T_1)))| =$$

$$= |\mathbf{h}_\infty(\operatorname{ct}(T_k))| + |\mathbf{h}_\infty(\operatorname{ct}(T_1))| \le 2A.$$

Since $O_{\tau(R)}(\mathbf{r})$ is the maximum of all the sums of the type $|\mathbf{r}_1 + ... + \mathbf{r}_{k-1}|$ the lemma follows. □

Next, we determine a set of independent generators for the homology group $H_1(\overline{S}, \operatorname{Cusp}(S))$. Let $\mathcal{D}$ be an ideal fundamental polygon for G so that the edges of $\mathcal{D}$ are geodesics from $\lambda(G)$ (every such $\mathcal{D}$ has $2(2g + n - 1)$ edges). Choose an orientation on every such edge, so that the identification (given by G) respects the orientation. Denote by $\lambda^*_{\mathrm{Gen}}(S) \subset \lambda^*(S)$ the set of oriented geodesics, whose lifts to the universal cover $\mathbf{H}$ are the oriented geodesics from the boundary of $\mathcal{D}$ (if we forget about the orientation, the corresponding set of geodesic is called $\lambda_{\mathrm{Gen}}(S)$). The set $\lambda^*_{\mathrm{Gen}}(S)$ contains $2g + n - 1$ different elements. Note that the complement of the set $\lambda^*_{\mathrm{Gen}}(S)$ in S is simply connected. This shows that the geodesics from $\lambda^*_{\mathrm{Gen}}(S)$ generate the group $H_1(\overline{S}, \operatorname{Cusp}(S))$. On the other hand, the dimension of the group is $H_1(\overline{S}, \mathrm{Cusp}(S))$ is $2g+n-1$. This shows that the set $\lambda^*_{\mathrm{Gen}}(S)$ is a set of independent generators. The following proposition is elementary.

Proposition 3.2. *Let $R \in \mathbf{N}\mathcal{T}(S)$. Assume that*

$$\partial R = \sum_{i=1}^{i=m} k_i \gamma^*_{k_i},$$

where $\gamma^*_{k_i} \in \lambda^*_{\mathrm{Gen}}(S)$ and $k_i \in \mathbf{Z}$. Then $k_i = 0$ for every i that is $\partial R = 0$ in $\mathbf{Z}\Gamma^*(S)$.

Proof. We have already observed that the homology class of ∂R is equal to zero in $H_1(\overline{S}, \mathrm{Cusp}(S))$. Since $\lambda^*_{\mathrm{Gen}}(S)$ are independent generators the proposition follows. □

3.3. **The Correction lemma.** Given a labelled collection of triangles $\mathcal{C}$ our aim is to equip $\mathcal{C}$ with the corresponding involution to produce an admissible pair. This of course not always possible. As we saw above, only if $\partial\mathcal{C} = 0$ in $\Gamma^*(S)$ one can do this. Moreover, we want to be able to construct this involution so that the centres of the triangles that are paired are as close as possible. Given an arbitrary $\mathcal{C}$ we can add more labels to the set $\mathcal{L}_{\mathcal{C}}$ and expand the domain and the range of the map $\mathrm{lab}_{\mathcal{C}}$ so that we are able to construct the corresponding involution. The purpose of this subsection is to prove the lemma which tells us how many labels we need to add.

We first prove a few propositions. All the distances mentioned below are considered to in the hyperbolic metric (unless stated otherwise).

Proposition 3.3. *There exists a constant $d_1(S) > 0$ which depends only on S so that the following holds. Let $c \in \mathrm{Cusp}$ and let $G = G_c$ $c = \infty$ be normalised. Let $\gamma_0 \in \Gamma(G)$ with the endpoints $c_1, c_2 \in \mathrm{Cusp}(G)$, $c_1, c_2 \in \mathbf{R}$, $c_1 < c_2$, and let $z_0 \in \gamma$. Assume that for $\infty \in \mathrm{Cusp}(G)$ the geodesic γ_0 intersects at least $100N(S)$ geodesics from $\lambda(S)$ that end at ∞. Also, assume that $z_0 \in \mathcal{H}_\infty(t_2(S))$. Let γ' be the geodesic that is orthogonal to γ_0 and that contains z_0. Let $c_1 < x < c_2$ and $y \in \mathbf{R}$ be the endpoints of γ'. We have:*

- *There exists $c_3 \in \mathrm{Cusp}(G)$ so that the geodesic that connects c_3 and ∞ belongs to $\lambda(G)$ and so that c_3 is the closest point to x subject to the condition $c_1 + 2 < c_3 < c_2 - 2$. Let γ_i, $i = 1, 2$, be the geodesic that connects c_i with c_3 and let T be the triangle bounded by γ_0, γ_1 and γ_2. Then the distance between the centre $\mathrm{ct}(T)$ and the point z_0 is less than $d_1(S)$.*
- *Assume $y > c_2$ (the analogous statement holds for $y < c_1$). There exists $c_3 \in \mathrm{Cusp}(G)$ so that the geodesic that connects c_3 and ∞ belongs to $\lambda(G)$ and so that c_3 is the closest point to y subject to the condition $c_2 + 2 < c_3$. Let γ_i, $i = 1, 2$, be the geodesic that connects c_i with c_3 and let T be the triangle bounded by γ_0, γ_1 and γ_2. Then the distance between the centre $\mathrm{ct}(T)$, and the point z_0 is less than $d_1(S)$.*

Proof. The existence of such c_3 in both cases follows from the fact that $|c_1 - c_2| \geq 100$. We prove the first statement. Assume that $|c_3 - c_2| \leq |c_3 - c_1|$ (the other case is done in the same way). Clearly

$$|x - c_3| < 3. \tag{19}$$

Since $z \in \mathcal{H}_\infty(t_2(S))$ it is elementary to see that

$$c_2 - x > \frac{e^{t_2(S)}}{2}. \tag{20}$$

Remark. Let $z_0 = p+iq$. If $c_2 - x < 2$ then $c_2 - x < x - c_1$. This implies $x < p < c_2$. Then one shows that $\max\{(c_2 - p), (p - x)\} \geq \frac{q}{2} \geq \frac{e^{t_2(S)}}{2}$.

We prove the existence of the constant $d_1(S)$ by contradiction. Assume that there is no such $d_1(S)$. Then there exists the corresponding sequence of geodesic $\gamma_0(m)$ and points $z_0(m)$, $m \in \mathbf{N}$, so that $\mathbf{d}(\operatorname{ct}(T(m)), z_0(m)) \to \infty$ when $m \to \infty$ (here $T(m)$ is the corresponding sequence of triangles). Let

$$f_m(z) = \frac{(z - c_2(m))}{|c_3(m) - c_2(m)|},$$

where $c_1(m)$, $c_2(m)$, $c_3(m)$ and $x(m)$, are the corresponding sequences of points. Since $|c_3(m) - c_2(m)| \leq |c_3(m) - c_1(m)|$, we have that the sequence of triangles $f_m(T(m))$ after passing onto a subsequence if necessary, converges to a non-degenerate ideal triangle $T(\infty)$ in $\mathbf{H}$. From the choice of f_m we have that $f_m(c_2(m)) = 0$ for every $m \in \mathbf{N}$.

If $|c_3(m) - c_2(m)| \to \infty$ then from (19) we conclude that the sequence $f_m(x(m))$ converges to $x(\infty) < 0$. If $|c_3(m) - c_2(m)|$ remains bounded then from (20), and from the fact that $|c_2 - c_3| \geq 2$ (which means that this sequence is bounded from below as well), we conclude that the sequence $f_m(x(m))$ converges to $x(\infty) < 0$. Also, $f_m(c_1(m)) \to c_1(\infty)$ and $c_1(\infty) < x(\infty)$. This shows that the sequence of geodesic $f_m(\gamma_0(m))$ converges to a proper geodesic in $\mathbf{H}$. Moreover, $f_m(z_0(m))$ converges to a point $z_0(\infty) \in \mathbf{H}$. Since $\mathbf{d}(\operatorname{ct}(T(m)), z_0(m)) \to \mathbf{d}(\operatorname{ct}(T(\infty)), z_0(\infty))$, $m \to \infty$, we obtain a contradiction. The second statement is proved in a similar way. □

Proposition 3.4. *There exist constants $d_2(S), d_3(S), \theta(S) > 0$ so that the following holds. Let $\gamma_0 \in \Gamma(G)$ and suppose $\iota(\gamma_0, \tau(G)) > 0$. Let $z_0 \in \gamma_0 \cap \mathbf{Th}_G(t_1(S))$. Let $\gamma_1 \in \lambda(G)$ be the geodesic, so that letting $w = \gamma_0 \cap \gamma_1$ the distance $\mathbf{d}(z_0, w)$ is the smallest among all the distances between z_0 and the intersection points between γ_0 and geodesics from $\lambda(G)$ (if there are two such closest points w we choose either one of them). Assume in addition that if for some $c \in \operatorname{Cusp}(G)$ the geodesic γ_0 intersects at least $100N(S)$ geodesics from $\lambda(G)$ that all have c as their endpoint, then z_0 does not belong to the horoball $\mathcal{H}_c(t_2(S) - 1)$. Then,*

1. *We have $\mathbf{d}(z_0, w) < d_2(S)$ and the smaller angle between γ_0 and γ_1 is greater than $\theta(S)$.*
2. *Let $c_1 \in \operatorname{Cusp}(G)$ be an endpoint of γ_1. Let $\gamma_0', \gamma_0'' \in \Gamma(G)$ be the geodesics that connect c_1 with the two endpoints of γ_0 respectively. Let $T \in \mathcal{T}(G)$ be the triangle bounded by γ_0 γ_0' and γ_0''. Then $\iota(\gamma_0', \tau(G)) + \iota(\gamma_0'', \tau(G)) \leq \iota(\gamma_0, \tau(G)) - 1$. Moreover, there are points $z_0' \in \gamma_0' \cap \mathbf{Th}_G(t_1(S))$ and $z_0'' \in \gamma_0'' \cap \mathbf{Th}_G(t_1(S))$ so that the distance between the centre of the triangle T and any of the points z_0, z_0', z_0'' is less than $d_3(S)$.*

Proof. We first prove (1). Let $\gamma_0(m)$, $m \in \mathbf{N}$, be a sequence of geodesics, and $z_0(m) \in \gamma_0(m)$ a sequence of points, so that $\gamma_0(m)$ and $z_0(m)$ share the above stated properties of γ_0 and z_0. Let $\gamma_1(m) \in \tau(G)$ and $w(m) \in \gamma_1(m)$ be the corresponding geodesics and points. Assume that either $\mathbf{d}(z_0(m), w(m)) \to \infty$ or that the angle between $\gamma_0(m)$ and $\gamma_1(m)$ tends to zero, when $m \to \infty$. Since $z_0(m) \in \mathbf{Th}_G(t_1(S))$ we can choose $f_m \in G$ so that after passing to a subsequence if necessary, we have $f_m(z_0(m)) \to z_0(\infty) \in \mathbf{H}$. Therefore, we may assume that $z_0(m) \to z_0(\infty) \in \mathbf{Th}_G(t_1(S))$. Let $\gamma_0(\infty)$ be a geodesic in $\mathbf{H}$ where $\gamma_0(m) \to \gamma_0(\infty)$. If $\mathbf{d}(z_0(m), w(m)) \to \infty$ then the geodesic $\gamma_0(\infty)$ does not intersect any geodesics from $\lambda(G)$. We conclude that $\gamma_0(\infty) \in \lambda(G)$. If the sequence $\{\mathbf{d}(z_0(m), w(m))\}$ is bounded, then the sequence of geodesics $\gamma_1(m)$ tends to a geodesic $\gamma_1(\infty) \in \lambda(G)$. If the angle between $\gamma_0(m)$ and $\gamma_1(m)$ tends to zero, when $m \to \infty$ then $\gamma_0(\infty) = \gamma_1(\infty)$ and we again conclude that $\gamma_0(\infty) \in \lambda(G)$.

Let $c_1(\infty), c_2(\infty) \in \text{Cusp}(G)$ denote the endpoints of $\gamma_0(\infty) \in \lambda(G)$. Since $z_0(\infty) \in \mathbf{Th}_G(t_1(S))$ we conclude from Proposition 3.1 that $z_0(\infty)$ belongs to both horoballs $\mathcal{H}_{c_1(\infty)}(t_2(S))$ and $\mathcal{H}_{c_2(\infty)}(t_2(S))$. Therefore, for m_0 large enough, and for every $m \ge m_0$ we have

$$z_0(m) \in \mathcal{H}_{c_1(\infty)}(t_2(S) - \frac{1}{2}),$$

(21)

$$z_0(m) \in \mathcal{H}_{c_2(\infty)}(t_2(S) - \frac{1}{2}).$$

Note that $\mathcal{H}_{c_1(\infty)}(t_2(S) - \frac{1}{2})$ is contained in the horoball $\mathcal{H}_{c_1(\infty)}(t_2(S) - 1)$ (the same is true for $c_2(\infty)$). On the other hand, since $\gamma_0(m) \to \gamma_0(\infty)$ and $\gamma_0(m) \neq \gamma_0(\infty)$ (since by the assumption of the lemma we have that $\gamma_0(m)$ does not belong to $\lambda(G)$), we conclude that the number of geodesics from $\lambda(G)$ that have either $c_1(\infty)$ or $c_2(\infty)$ as their endpoints, and that are intersected by $\gamma_0(m)$ tends to ∞. Let $m_1 > m_0$ be large enough, so that $\gamma_0(m_1)$ intersects at least $100N(S)+1$ geodesics from $\lambda(G)$ that all have $c_1(\infty)$ as their endpoint (the case when these geodesics end at $c_2(\infty)$ is similar). From this and from (21) we obtain a contradiction with the assumption that $z_0(m_1)$ does not belong to a $t_2(S) - 1$-horoball of any cusp $c \in \text{Cusp}(G)$ when $\gamma_0(m_1)$ intersects at least $100N(S)$ geodesics from $\lambda(G)$ that end at c.

Next we prove (2). Every geodesic from $\lambda(G)$ that intersects γ_0 either intersects exactly one of the geodesics γ_0' or. γ_0'' or this geodesic ends at c_1. Since $\gamma_1 \in \lambda(G)$ has c_1 as its endpoint, we see that $\iota(\gamma_0', \tau(G)) + \iota(\gamma_0'', \tau(G)) \le \iota(\gamma_0, \tau(G)) - 1$.

We prove the last part of (2) by contradiction. The argument is very similar as in the proof of (1). Let $\gamma_0(m) \in \Gamma(G)$, $m \in \mathbf{N}$, be a sequence of geodesics, and $z_0(m) \in \gamma_0(m) \cap \mathbf{Th}_G(t_1(S))$, $m \in \mathbf{N}$, a sequence of points. Let $\gamma_0'(m)$ and $\gamma_0''(m)$ be the corresponding sequences of geodesics, and let $T(m)$ be the corresponding sequence of triangles, where $\text{ct}(T(m))$ denotes the centres of $T(m)$. Assume that at least one of the distances $\mathbf{d}(\text{ct}(T(m)), z_0(m))$ $\mathbf{d}(\text{ct}(T(m)), \mathbf{Th}_G(t_1(S)) \cap \gamma_0'(m))$ or $\mathbf{d}(\text{ct}(T(m)), \mathbf{Th}_G(t_1(S)) \cap \gamma_0''(m))$ tends to ∞ when $m \to \infty$. Same as above, we can assume that $z_0(m) \to z_0(\infty) \in \mathbf{H}$ and $\gamma_0(m) \to \gamma_0(\infty)$. From the conclusion (1) of this lemma that was proved above, we conclude that $T(m)$ converges to an ideal triangle $T(\infty)$. In particular, $\gamma_0'(m) \to \gamma_0'(\infty)$ and $\gamma_0''(m) \to \gamma_0''(\infty)$. Also, $\text{ct}(T(m)) \to \text{ct}(T(\infty))$ where $\text{ct}(T(\infty))$ is the centre of $T(\infty)$. Since $\mathbf{d}(\text{ct}(T(m)), z_0(m)) \to \mathbf{d}(\text{ct}(T(\infty)), z_0(\infty))$ we see that $\mathbf{d}(\text{ct}(T(m)), z_0(m))$ does not tend to ∞. This means that one of the distances $\mathbf{d}(\text{ct}(T(m)), \mathbf{Th}_G(t_1(S)) \cap \gamma_0'(m))$ or $\mathbf{d}(\text{ct}(T(m)), \mathbf{Th}_G(t_1(S)) \cap \gamma_0''(m))$ tends to ∞ when $m \to \infty$. This implies that at least one of the geodesics $\gamma_0'(\infty)$ or $\gamma_0''(\infty)$ is contained in $\mathbf{H} \setminus \mathbf{Th}_G(t_1(S))$.

Since $t_1(S) > 0$ we have that the set $\mathbf{H} \setminus \mathbf{Th}_G(t_1(S))$ is a union of disjoint horoballs (the base points of these horoballs are in $\operatorname{Cusp}(G)$) in $\mathbf{H}$. Since every geodesic in $\mathbf{H}$ is a connected set, we conclude that at least one of the geodesics $\gamma_0'(\infty)$ or $\gamma_0''(\infty)$ is contained in a horoball in $\mathbf{H}$. This is a contradiction. □

We will use the following definition:

Definition 3.6. *Let $\gamma \in \lambda(G)$ be a geodesic. Let $T_1, T_2 \in \tau(G)$ be the two triangles that are adjacent along γ. By $\operatorname{mid}(\gamma)$ we denote the intersection point between γ and the geodesic segment that connects the centres of the triangles T_1 and T_2.*

We now prove the Correction Lemma.

Lemma 3.3. *Let $\gamma_0 \in \Gamma(G)$ be a geodesic, and let $\gamma_0^* \in \Gamma^*(G)$ be the geodesic γ_0 with a chosen orientation. Let $z_0 \in \gamma_0$ be a point. Then there exists an ideal polygon $\mathcal{P}$ with a triangulation $\tau(\mathcal{P})$ that has the following properties:*

1. *The geodesic γ_0 is an edge of $\mathcal{P}$ and all other edges of $\mathcal{P}$ are geodesics from $\lambda_{\text{Gen}}(G)$ where $\lambda_{\text{Gen}}(G)$ is the lift of $\lambda_{\text{Gen}}(S)$ to $\mathbf{H}$.*
2. *The polygon $\mathcal{P}$ is to the right of γ_0^*.*
3. *There exists a constant $D > 0$ which depends only on S so that the centres of any two adjacent triangles from $\mathcal{P}$ are within the D hyperbolic distance. If $T \in \tau(\mathcal{P})$ is the triangle adjacent to γ_0 then the distance between the centre of T and the point z_0 is less than $\frac{D}{2}$. If $T \in \tau(\mathcal{P})$ is a triangle adjacent to another edge γ of $\mathcal{P}$ then the hyperbolic distance between the centre of T and the point $\operatorname{mid}(\gamma) \in \gamma$ is less than $\frac{D}{2}$.*
4. *There exist constants $C, K > 0$ which depend only on S so that $|\tau(\mathcal{P})| \leq C(\mathcal{K}(\gamma_0, z_0)) + K$. Here $|\tau(\mathcal{P})|$ stands for the total number of triangles in $\mathcal{P}$.*

Proof. First consider the case when $\mathcal{K}(\gamma_0, z_0) = 0$. We allow any orientation on γ_0^*. Then we have $\gamma_0 \in \lambda(G)$. Moreover, let $c_1, c_2 \in \operatorname{Cusp}(G)$ be the two endpoint of γ_0. In this case we have that $\mathbf{h}_{c_1}(z_0), \mathbf{h}_{c_2}(z_0) \leq t_1(S)$. We build the corresponding polygon as follows. Let T be the triangle from $\tau(G)$ that is to the right of γ_0^*. If an edge of T belongs to the set $\lambda_{\text{Gen}}(G)$ we do not add any more triangles along this edge. If an edge of T does not belong to $\lambda_{\text{Gen}}(G)$ we glue the corresponding triangle from $\tau(G)$ to the right of this edge. We repeat this process until all the edges of the polygon $\mathcal{P}$ we obtained are all from $\lambda_{\text{Gen}}(G)$ (except the edge γ_0). This process has to end because the geodesics from $\lambda_{\text{Gen}}(G)$ are the edges of an ideal fundamental polygon for G. This way we have also constructed the corresponding triangulation $\tau(\mathcal{P})$ of the polygon $\mathcal{P}$.

We repeat this construction for every $\gamma_0 \in \lambda(G)$. If $f \in G$ and $f(\gamma_0) = \gamma_0'$ then the corresponding polygon for γ_0' is the image of the corresponding polygon for γ_0 under the map f. So, we need to consider only $|\lambda(S)|$ different geodesics from $\lambda(G)$. Since $\gamma_0 \cap \mathbf{Th}_G(t_1(S))$ is a compact set in $\mathbf{H}$ by the compactness argument we find that there exist constants $D_1 > 0$ and $K > 0$ so that for every such $\gamma_0 \in \lambda(G)$ the corresponding polygon $\mathcal{P}$ satisfies that $|\tau(\mathcal{P})| \leq K$. Also, the distances between the centres of adjacent triangles in $\mathcal{P}$ are bounded by D_1. The distances between the centres of boundary triangles, and the corresponding points on the edges of $\mathcal{P}$ are bounded by $\frac{D_1}{2}$.

We now prove the general case by induction. We first determine the constants D, C from the statement of the lemma (the constant K has already been defined).

Set

$$D = 2\big(D_1 + 2d_3(S) + (64 + \log 32) + 2d_1(S) + 2\log(4N(S)+1)\big)$$

Let

$$C_1 = \frac{K+1}{\frac{[e^{t_1(S)}]}{2} - 2 - N(S)}.$$

Recall that we chose $t_1(S)$ large enough so that the denominator is a positive constant. Set

$$C = (2K+1) + 10K + C_1.$$

The numbers that appear in the definition of C and D, will appear in the arguments below.

The proof is by induction on $m = \mathcal{K}(\gamma_0, z_0)$. The case $0 = m = \mathcal{K}(\gamma_0, z_0)$ has been done above. Fix $m \in \mathbf{N}$ and assume that the statement of the lemma is true for every non-negative integer that is less than m. We now prove that the statement for m. There are four cases to consider.

Case 1. In this case we assume that $z_0 \in \mathbf{Th}_G(t_1(S))$. Moreover, we assume that if for some $c \in \operatorname{Cusp}(G)$ γ_0 intersects at least $100N(S)$ geodesics from $\lambda(G)$ that all end at c then the point z_0 does not belong to $\mathcal{H}_c(t_2(S))$ (then z_0 does not belong to $\mathcal{H}_c(t_2(S)-1)$ either, so we may apply Proposition 3.4). We allow any orientation γ_0^*. We have $\mathcal{K}(\gamma_0, z_0) = \iota(\gamma_0, \tau(G)) = m$.

We apply Proposition 3.4. That is, there are geodesics $\gamma_0', \gamma_0'' \in \Gamma(G)$ so that there exists a triangle $T \in \mathcal{T}(G)$ that is bounded by γ_0, γ_0' and γ_0''. Moreover, there are points $z_0' \in \gamma_0' \cap \mathbf{Th}_G(t_1(S))$ and $z_0'' \in \gamma_0'' \cap \mathbf{Th}_G(t_1(S))$ so that the distance between the centre of the triangle T and any of the points z_0, z_0', z_0'' is less than $d_3(S)$. Also, we have that $\iota(\gamma_0', \tau(G)) + \iota(\gamma_0'', \tau(G)) \le \iota(\gamma_0, \tau(G)) - 1 = m-1$. We choose the orientations of ${\gamma_0'}^*$ and ${\gamma_0''}^*$ so that γ_0 is to the left of both of them.

Since $z_0', z_0'' \in \mathbf{Th}_G(t_1(S))$ we have $\mathcal{K}(\gamma_0', z_0) = \iota(\gamma_0', \tau(G))$ and $\mathcal{K}(\gamma_0'', z_0) = \iota(\gamma_0'', \tau(G))$. By the induction hypothesis, the statement is true for both pairs $({\gamma_0'}^*, z_0')$ and $({\gamma_0''}^*, z_0'')$. We glue the corresponding triangulated polygons to the right of ${\gamma_0'}^*$ and ${\gamma_0''}^*$ respectively. Together with the triangle T this gives the needed polygon $\mathcal{P}$. We have (using $C > K+1$)

$$|\tau(\mathcal{P})| \le 1 + C\mathcal{K}(\gamma_0', z_0') + C\mathcal{K}(\gamma_0'', z_0'') + 2K = (2K+1) + C(\mathcal{K}(\gamma_0, z_0) - 1) \le C\mathcal{K}(\gamma_0, z_0) + K.$$

By the induction hypothesis, the distances between the centres of adjacent triangles in the two polygons we have glued along γ_0' and γ_0'' are less than D. Also, by the induction hypothesis, the distances between the points z_0' and z_0'' and the centres of the corresponding triangles in those glued polygons, that contain the points z_0' and z_0'' in the their boundaries respectively, are less than $\frac{D}{2}$. Since the distance between the centre of the triangle T and any of the points z_0, z_0', z_0'' is less than $d_3(S) < \frac{D}{2}$ we have that the distance between the centre of T and either one of the two adjacent triangles to T in $\mathcal{P}$ is less than D. This shows that all the distances between the corresponding points in $\mathcal{P}$ are within D or $\frac{D}{2}$ as required. This settles the first case.

Case 2. Let $c_1, c_2 \in \operatorname{Cusp}(G)$ be the endpoints of γ_0. In this case we assume that $\max\{\mathbf{h}_{c_1}(z_0), \mathbf{h}_{c_2}(z_0)\} > t_1(S)$. For the sake of the argument assume $\mathbf{h}_{c_1}(z_0) > t_1(S)$. We allow any orientation on γ_0. Normalise G that is $G = G_{c_1}$ and $c_1 = \infty$. Also, assume that the orientation of γ_0^* is such that the pair (c_2, ∞) has the positive orientation on γ_0^* (the other case is done analogously). Let $k = \iota(\gamma_0, z_0)$. Then

$[e^{\mathbf{h}_\infty(z_0)}] = m - k$. Let $c_3 \in \mathrm{Cusp}(G)$ be the point on $\mathbf{R}$ defined as follows. We require that the geodesic that connects c_3 and ∞ belongs to $\lambda(G)$. Then, c_3 is the smallest such point subject to the condition

$$c_2 + \frac{e^{\mathbf{h}_\infty(z_0)}}{4N(S)} \le c_3.$$

Let $\gamma_1 \in \Gamma(G)$ be the geodesic that connects c_2 and c_3 (we fix the orientation of γ_1^* so that ∞ is to the left of γ_1^*). Since G is normalised, we have

$$\frac{e^{\mathbf{h}_\infty(z_0)}}{4N(S)} < c_3 - c_2 < \frac{e^{\mathbf{h}_\infty(z_0)}}{4N(S)} + 1,$$

and this yields

$$\iota(\gamma_1, \tau(G)) \le k + \left(\frac{e^{\mathbf{h}_\infty(z_0)}}{4N(S)} + 1 \right) N_\infty(S) \le k + \frac{[e^{\mathbf{h}_\infty(z_0)}]}{4} + 1 + N(S).$$

Let $z_1 = \mathbf{z}_{\max}(\gamma_1, \infty)$. Since $(c_3 - c_2) \ge \frac{1}{4N(S)}$ we have that $z_1 \in \mathcal{H}_\infty(-\log(8N(S)))$ and therefore by the choice of $t_1(S)$ in Proposition 3.1 the point z_1 does not belong to any $t_1(S)$-horoball (except possibly the one at ∞ which is not an endpoint of γ_1). This implies that

$$\mathcal{K}(\gamma_1, z_1) = \iota(\gamma_1, \tau(G)).$$

Let γ_2 be the geodesic that connects c_3 and ∞ (we fix the orientation of γ_2^* so that c_2 is to the left of γ_2^*). By the choice of c_3 we have that $\gamma_2 \in \lambda(G)$. Let $z_2 \in \gamma_2$ be the point whose imaginary part is $\frac{e^{\mathbf{h}_\infty(z_0)}}{4}$. Since

$$\frac{e^{\mathbf{h}_\infty(z_0)}}{4} > \frac{e^{t_1(S)}}{4} > 1,$$

we have that z_2 does not belong to any $t_1(S)$-horoball except possibly the one at ∞. We have

$$\mathcal{K}(\gamma_2, z_2) \le [e^{\mathbf{h}_\infty(z_2)}] \le \frac{e^{\mathbf{h}_\infty(z_0)}}{4} \le \frac{[e^{\mathbf{h}_\infty(z_0)}]}{4} + 1.$$

By the induction hypothesis, we have that the statement of the lemma is true for both pairs (γ_1^*, z_1) and (γ_2^*, z_2). We construct the polygon $\mathcal{P}$ by gluing the corresponding polygon for γ_1 to the right of γ_1^* and the corresponding polygon for γ_2 to the right of γ_2^*. We also add the triangle $T \in \mathcal{T}(G)$ bounded by γ_0 γ_1 and γ_2 to $\mathcal{P}$. Combining this with the above estimates (and the definition of C) we have

$$|\mathcal{P}| \le 1 + C(k + \frac{[e^{\mathbf{h}_\infty(z_0)}]}{4} + 1 + N(S)) + C(\frac{[e^{\mathbf{h}_\infty(z_0)}]}{4} + 1) + 2K \le$$

$$\le 1 + C(k + \frac{[e^{\mathbf{h}_\infty(z_0)}]}{2} + 2 + N(S)) + 2K \le$$

$$\le C(k + [e^{\mathbf{h}_\infty(z_0)}]) + K + \big(K + 1 - C(\frac{[e^{t_1(S)}]}{2} - 2 - N(S))\big) \le C\mathcal{K}(\gamma_0, z_0) + K$$

Note that the distance between the centre of T and any of the points z_0, z_1, z_2 is less than $\log(4N(S)+1)$. One now shows that the distances between the corresponding points in $\mathcal{P}$ are within D or $\frac{D}{2}$ in much the same way as in the Case 1.

Case 3. The only case left to consider is when $z_0 \in \mathbf{Th}_G(t_1(S))$ but when there exists $c \in \mathrm{Cusp}(G)$ so that z_0 belongs to $\mathcal{H}_c(t_2(S))$ and where γ_0 intersects at least $100N(S)$ geodesics from $\lambda(G)$ that end at c. In this case γ_0 can not have c as its endpoint. Moreover, $\mathcal{K}(\gamma_0, z_0) = \iota(\gamma_0, \tau(G)) = m$. Assume that G is normalised, $G = G_c$. Let $c_1, c_2 \in \operatorname{Cusp}(G)$, $c_1 < c_2$, be the endpoints of γ_0. We first consider the case when the orientation of γ_0^* is such that the cusp c is to the left of γ_0.

Let γ_0' be the geodesic that is orthogonal to γ_0 and that contains z_0. Let $c_1 < x < c_2$ be the corresponding endpoint of γ_0'. Since γ_0 intersects at least $100N(S)$ geodesics from $\lambda(G)$ that end at ∞ we can choose a point $c_3 \in \operatorname{Cusp}(G)$ so that the geodesic that connects c_3 and ∞ belongs to $\lambda(G)$ and so that c_3 is the closest such point to the point x subject to the condition $c_1 + 2 < c_3 < c_2 - 2$ (see Proposition 3.3). Let $\gamma_i \in \Gamma(G)$ $i = 1, 2$, be the geodesic that connects c_i with c_3. Choose the orientation of γ_i^* so that ∞ is to the left of γ_i^*. Let $T \in \mathcal{T}(G)$ be the triangle bounded by γ_0, γ_1 and γ_2. Let z_i, $i = 1, 2$, be the orthogonal projections of the centre $\operatorname{ct}(T)$ to the geodesic γ_i. Since $|c_3 - c_i| > 2$, $i = 1, 2$, it is easily seen that $\mathbf{h}_\infty(z_i) > 0$. We conclude that the point z_i does not belong to the $t_1(S)$-horoballs at c_i or c_3. Therefore,

$$\mathcal{K}(\gamma_i, z_i) = \iota(\gamma_i, \tau(G)), i = 1, 2.$$

From Proposition 3.3 we have that $\iota(\gamma_1, \tau(G)) + \iota(\gamma_2, \tau(G)) = \iota(\gamma_0, \tau(G)) - 1 = \mathcal{K}(\gamma_0, z_0) - 1$. By the induction hypothesis, the statement of the lemma is true for the pair (γ_i, z_i). We glue the corresponding polygons to the right of γ_i^*. Let $T \in \mathcal{T}(G)$ be the triangle bounded by γ_0, γ_1 and γ_2. We add this triangle to obtain the required polygon $\mathcal{P}$. We have

$$|\tau(\mathcal{P})| \leq 1 + C\mathcal{K}(\gamma_1, z_1) + K + C\mathcal{K}(\gamma_2, z_2) + K \leq C\mathcal{K}(\gamma_0, z_0) + (2K+1) - C \leq C\mathcal{K}(\gamma_0, z_0) + K.$$

By Proposition 3.3 the distance between the centre of T and the point z_0 is less than $d_1(S)$. By the definition of z_1 and z_2 the distance between the centre of T and the point z_i is less than 1 (the distance between the centre of an ideal triangle, and its orthogonal projections to one of its sides is less than 1). In much the same way as before one shows that the distances between the corresponding points in $\mathcal{P}$ are within the required bounds.

Case 4. The assumptions are the same as in Case 3. except that the orientation of γ_0^* is such that ∞ is to the right of γ_0^*. Let y denote the second endpoint of γ_0' and assume that $y > c_2$ (the case $y < c_1$ is treated similarly). Recall $\mathcal{K}(\gamma_0, z_0) = \iota(\gamma_0, \tau(G)) = m$. Let m_3 denote the number of geodesics from $\lambda(G)$ that all end at ∞ and that are intersected by γ_0. Let m_1 denote the number number of geodesics from $\lambda(G)$ that are intersected by γ_0 and that are to the left of those geodesics we counted for m_3. Let m_2 denote the number of geodesics from $\lambda(G)$ that are intersected by γ_0 and that are to the right of those counted for m_3. Then $m_1 + m_2 + m_3 = m$.

Choose a point $c_3 \in \operatorname{Cusp}(G)$ (see Figure 3.3) so that the geodesic that connects c_3 and ∞ belongs to $\lambda(G)$ and so that c_3 is the closest such point to the point y subject to the condition $c_2 + 2 < c_3$ (see Proposition 3.3). Let k be the integer so that

$$\log_2[c_2 - c_1] - \log_2[c_3 - c_2] < k \leq \log_2[c_2 - c_1] - \log_2[c_3 - c_2] + 1, \tag{22}$$

where $[c_2 - c_1]$ and $[c_3 - c_2]$ denote the corresponding integer parts. Note that $m_3 \geq [c_2 - c_1]N_\infty(S) - 1$. This gives

$$\mathcal{K}(\gamma_0, z_0) \geq [c_2 - c_1]N_\infty(S).$$

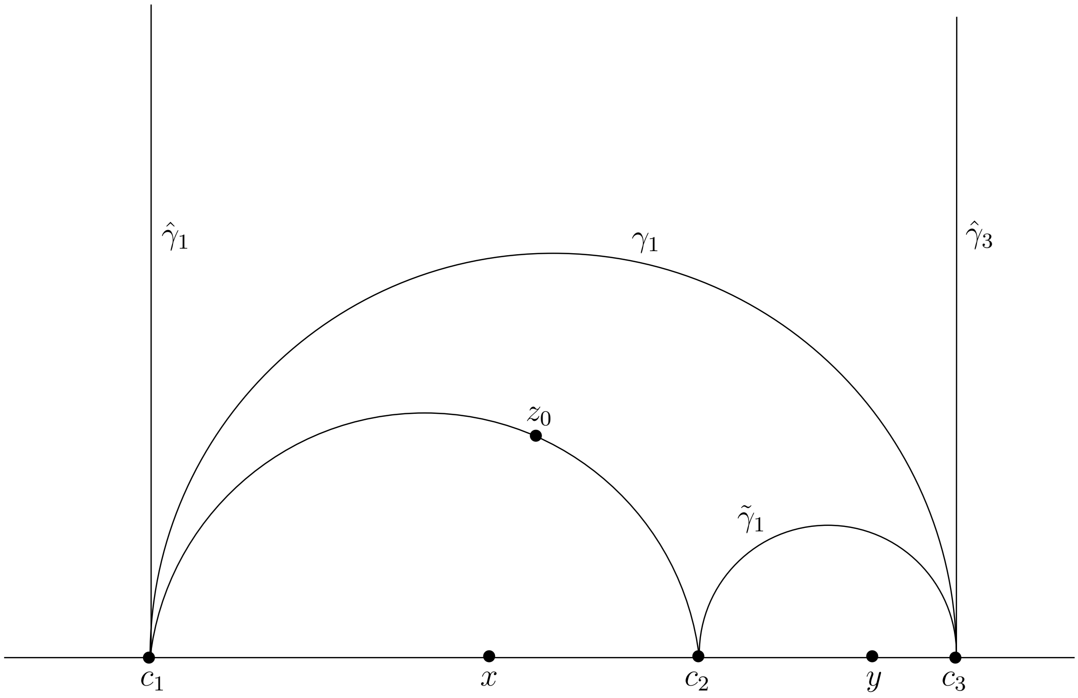

FIGURE 3

First consider the case $k > 3$. Let c_{i+3}, $i = 1, ..., k-3$, be given by

$$c_{i+3} = c_3 + \sum_{j=i}^{j=i} 2^j [c_3 - c_2].$$

Let $\gamma_i \in \Gamma(G)$, $i = 0, ..., k-2$, denote the geodesic that connects c_1 and c_{i+2}. By $\widetilde{\gamma}_i$ denote the geodesic that connects c_{i+1} and c_{i+2}, $i = 0, ..., k-2$ (with this definition, the geodesics γ_0 and $\widetilde{\gamma}_0$ agree). Let $\widehat{\gamma}_1$ be the geodesic that connects c_1 and ∞, and $\widehat{\gamma}_k$ the one that connects c_k and ∞. Let T_{i+1} be the triangle bounded by γ_i, γ_{i+1} and $\widetilde{\gamma}_{i+1}$. By $\widehat{T}$ denote the triangle bounded by $\widehat{\gamma}_1$, $\widehat{\gamma}_k$ and γ_{k-2}.

Note that by the construction, the distance between the centres of adjacent triangles T_i and T_{i+1} is less than $\log 2 < 1$. It is directly seen that the distance between $\mathrm{ct}(T_{k-2})$ and $\mathrm{ct}(\widehat{T})$ is less than $\log 32 + 64$. By Proposition 3.3 the distance between z_0 and the centre of the triangle T_1 is less than $d_1(S)$.

Let $\widetilde{z}_i \in \widetilde{\gamma}_i$, $i = 1, ..., k-2$, be the points that are the orthogonal projections of the centres of triangles T_i to $\widetilde{\gamma}_i$. Since $(c_3 - c_2) > 2$ we have that the centre of T_i is high enough with respect to ∞ so that $\mathbf{h}_\infty(\widetilde{z}_i) > 0$. Therefore, $\mathcal{K}(\widetilde{\gamma}_i, \widetilde{z}_i) = \iota(\widetilde{\gamma}_i, \tau(G))$. Since for both endpoints of γ_i, $i \geq 2$, we have that there are geodesics from $\lambda(G)$ that connect those points to ∞ we have that

$$\mathcal{K}(\widetilde{\gamma}_i, \widetilde{z}_i) \leq 2^i [c_3 - c_2] N_\infty(S) - 1 + m_1 + m_2,$$

for $i \geq 2$. For the pair $(\widetilde{\gamma}_1, \widetilde{z}_1)$ we have

$$\mathcal{K}(\widetilde{\gamma}_1, \widetilde{z}_1) \leq m_2 + 2[c_3 - c_2]N_\infty(S).$$

On the other hand, let $\widehat{z}_1 \in \widehat{\gamma}_1$ and $\widehat{z}_k \in \widehat{\gamma}_k$ be the points whose imaginary parts are equal to $\frac{[c_2-c_1]}{8}$. Then

$$\mathcal{K}(\widehat{\gamma}_1, \widehat{z}_1) \leq m_1 + \frac{[c_2 - c_1]}{8},$$

and

$$\mathcal{K}(\widehat{\gamma}_k, \widehat{z}_k) \leq \frac{[c_2 - c_1]}{8}.$$

Also, it is easily seen that the distances between the centre of the triangle $\widehat{T}$ and the points $\widehat{z}_1$ and $\widehat{z}_k$ are less than $\log 64 < 8$.

We now apply the induction hypothesis on the pairs $(\widetilde{\gamma}_i, \widetilde{z}_i)$, $i = 1, ..., k-2$, $(\widehat{\gamma}_1, \widehat{z}_1)$, $(\widehat{\gamma}_{k-2}, \widehat{z}_{k-2})$. We put the corresponding orientations on these geodesics, so that γ_0 is to the left of them. After gluing the corresponding polygons to the right of the corresponding geodesics, and adding the triangles T_i and $\widehat{T}$ we obtain the polygon $\mathcal{P}$. We have

$$|\tau(\mathcal{P})| \leq (k-2)+1+C\sum_{i=1}^{i=k-2} 2^i[c_3-c_2]N_\infty(S)+Cm_2+C(m_1+\frac{[c_2-c_1]}{8})+C\frac{[c_2-c_1]}{8}+kK \leq$$

$$\leq (k+1)K - 1 + C\big(m_1 + m_2 + \frac{[c_2 - c_1]}{4} + 2^{k-1}[c_3 - c_2]N_\infty(S)\big) \leq$$

$$\leq (k+1)K - 1 + C\big(m_1 + m_2 + \frac{[c_2 - c_1]}{4} + \frac{[c_2 - c_1]}{2}N_\infty(S) + \frac{1}{2}\big) \leq$$

$$\leq C(m_1+m_2+[c_2-c_1]N_\infty(S)-1)+K+\left(\frac{3}{2}C - C\frac{[c_2 - c_1]}{4} + kK\right) \leq C\mathcal{K}(\gamma_0, z_0)+K.$$

The last part of the above estimate follows from (22), the fact that $c_2 - c_1 > 100$ and the property $C > 10K$.

Let $0 \leq k \leq 3$ in the above construction. Let $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ be the geodesics that connect c_1 and c_2 with ∞, respectively. Let $\widehat{T}$ be the triangle bounded by γ_0, $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$. Let $\widehat{z}_1 \in \widehat{\gamma}_1$ and $\widehat{z}_2 \in \widehat{\gamma}_2$ be the corresponding points (chosen in the same way as above). We now glue the corresponding polygons (whose existence follows from the induction hypothesis), to the appropriate sides of $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$. We also add the triangle $\widehat{T}$ to obtain the polygon $\mathcal{P}$. The rest is proved in the same way as in the case $k > 3$. □

4. Measures on triangles and the proof of Theorem 1.1

4.1. **Measures on triangles and the $\widehat{\partial}$ operator.** Let $\mathbf{N}^1\Gamma(S)$ and $\mathbf{N}^1\Gamma(G)$ denote the unit normal bundles of $\Gamma(S)$ and $\Gamma(G)$ respectively. Here $\mathbf{N}^1\Gamma(S)$ is the union of the normal bundles $N^1\gamma$, where $\gamma \in \Gamma(S)$. Every point in $N^1\gamma$ is a pair $(z, \vec{n})$ where $z \in \gamma$ and $\vec{n}$ is the unit vector in the tangent space T^1S at z that is orthogonal to γ. Every point $(z, \vec{n}) \in N^1\gamma$ corresponds to a unique pair (γ^*, z), where $\gamma^* \in \Gamma^*$ and $z \in \gamma$, such that the vector $\vec{n}$ points to the left of γ^*.

Connected components of $\mathbf{N}^1\Gamma(S)$ are the normal bundles $N^1\gamma^*$, where $\gamma^* \in \Gamma^*(S)$, and $\mathbf{N}^1\Gamma(S)$ is the disjoint union $\mathbf{N}^1\Gamma(S) = \cup_{\gamma^* \in \Gamma^*(S)} N^1\gamma^*$. The normal bundle $N^1\gamma^*$ is the subset of the normal bundle $N^1\gamma$ that contains the unit vectors that are pointing to the left of γ^*. Note that for a fixed geodesic γ and a chosen orientation γ^* the space $N^1\gamma^*$ is a connected 1-manifold. This manifold is identified with γ by the obvious projection map $N^1\gamma^* \to \gamma$. Therefore, the set $\mathbf{N}^1\Gamma(S)$ is a 1-manifold with countably many components. Note that the group G acts on $\mathbf{N}^1\Gamma(G)$ and $\mathbf{N}^1\Gamma(S) = \mathbf{N}^1\Gamma(G)/G$.

Remark. Since every point in $\mathbf{N}^1\Gamma(G)$ corresponds to a unique pair (γ^*, z). Therefore, we define the combinatorial length mapping $\mathcal{K} : \mathbf{N}^1\Gamma(G) \to \mathbf{N}$ as $\mathcal{K}(z, \vec{n}) = \mathcal{K}(\gamma^*, z)$, where $\vec{n}$ points to the left of γ^*. The induced map $\mathcal{K} : \mathbf{N}^1\Gamma(S) \to \mathbf{N}$ is also denoted by $\mathcal{K}$.

Definition 4.1. *Let γ be a geodesic in $\mathbf{H}$. We define the foot projection* $\operatorname{foot}_\gamma : \mathbf{H}\backslash\gamma \to N^1\gamma$ *as follows. Let $w \in \mathbf{H}\backslash\gamma$ and let $z \in \gamma$ denote the orthogonal projection of w to γ. Then $\operatorname{foot}_\gamma(w) \in N^1\gamma$ is the point $(z, \vec{n})$ such that the geodesic ray that starts at $(z, \vec{n})$ contains w.*

Unless otherwise stated, all measures in this paper are positive, Borel measures.

Definition 4.2. *If X is a topological space, then $\mathcal{M}(X)$ denotes the space of positive, Borel measures on X (necessarily finite). If X is a countable set, we equip X with the discrete topology. In particular, by $\mathcal{M}(\mathcal{T}(S))$ and $\mathcal{M}(\mathbf{N}^1\Gamma(S))$ we denote the spaces of positive, Borel measures on the set of triangles $\mathcal{T}(S)$ and the manifold $\mathbf{N}^1\Gamma(S)$. The corresponding spaces of measures on $\mathcal{T}(G)$ and $\mathbf{N}^1\Gamma(G)$ are denoted by $\mathcal{M}(\mathcal{T}(G))$ and $\mathcal{M}(\mathbf{N}^1\Gamma(G))$.*

Remark. Note that $\mathbf{N}\mathcal{T}(S)$ can be seen as the subset of $\mathcal{M}(\mathcal{T}(S))$. Each $R \in \mathbf{N}\mathcal{T}(S)$ induces a measure in $\mathcal{M}(\mathcal{T}(S))$ in the obvious way (if $R = k_1T_1 + ... + k_mT_m$ then the corresponding measure $\mu \in \mathcal{M}(\mathcal{T}(S))$ satisfies that $\mu(T_i) = k_i$).

We define the $\widehat{\partial} : \mathcal{M}(\mathcal{T}(S)) \to \mathcal{M}(\mathbf{N}^1\Gamma(S))$ operator as follows. The set $\mathcal{T}(S)$ is a countable set, so every measure from $\mathcal{M}(\mathcal{T}(S))$ is determined by its value on every triangle in $\mathcal{T}(S)$. Let $T \in \mathcal{T}(S)$ and let $\gamma_i \in \Gamma(S)$, $i = 1, 2, 3$, denote its edges. Set $V_i = \operatorname{foot}_{\gamma_i}(\operatorname{ct}(T))$ where $\operatorname{ct}(T)$ is the centre of T. Choose $\mu \in \mathcal{M}(\mathcal{T}(S))$. Let $\alpha^T \in \mathcal{M}(\mathbf{N}^1\Gamma(S))$ be the atomic measure, supported on V_1, V_2 and V_3 such that $\alpha^T(V_i) = \mu(T)$. We define $\widehat{\partial}\mu = \alpha \in \mathcal{M}(\mathbf{N}^1\Gamma(S))$ as

$$\alpha = \sum_{T \in \mathcal{T}(S)} \alpha^T.$$

If $\mu \in \mathcal{M}(\mathcal{T}(S))$ is a finite measure, then $\widehat{\partial}\mu$ is a finite measure as well. Moreover, the total measure of $\widehat{\partial}\mu$ is three times the total measure of μ.

4.2. **Transport of measure.** In this subsection we define the notion of equivalent measures. The following is the standard result in measure theory.

Proposition 4.1. *Let $\mu_i \in \mathcal{M}(\mathbf{R})$, $i = 1, 2$, be two finite measures on the real line $\mathbf{R}$. Let $K > 0$. The following conditions are equivalent:*

(1) *For every $a \in \mathbf{R}$ the following inequalities hold*

$$\mu_1(-\infty, a] \le \mu_2(-\infty, a + K],\ \mu_2(-\infty, a] \le \mu_1(-\infty, a + K]. \tag{23}$$

(2) *The total μ_1 and μ_2 measures of $\mathbf{R}$ coincide, that is $\mu_1(\mathbf{R}) = \mu_2(\mathbf{R})$. There exist mappings $\psi_i : (0, \mu_1(\mathbf{R})] \to \mathbf{R}$ so that $(\psi_i)_*(\nu) = \mu_i$ and $||\psi_1 - \psi_2||_\infty \le$*

K. Here ν denotes the Lebesgue measure on the interval $(0, \mu_1(\mathbf{R})]$ and $(\psi_i)_*(\nu)$ denotes the push-forward of the measure ν under ψ_i. Moreover, the mapping ψ_i is non-decreasing and left continuous.

(3) *There exists a topological space X with a Borel measure $\eta \in \mathcal{M}(X)$ so that the following holds. There exist mappings $\psi_i : X \to \mathbf{R}$ so that $(\psi_i)_*(\eta) = \mu_i$ and $||\psi_1 - \psi_2||_\infty \leq K$.*

If either of these three conditions is satisfied we say that μ_1 and μ_2 are K-equivalent measures.

Remark. If $g(x) = x + c$ is a translation, then the measures μ_1 and μ_2 are K-equivalent, if and only if the measures $g_*\mu_1$ and $g_*\mu_2$ are. Here $g_*\mu_i$ denotes the push-forward of the measure μ_i by g. Since g is a homeomorphism the similar statement is true for the pull-backs of μ_1 and μ_2. Also, note that the relation of being K-equivalent is symmetric. Moreover, if a pair of measures μ_1, μ_2 is K_1 equivalent, and a pair ν_1, ν_2 is K_2-equivalent, then the measures $\mu_1 + \nu_1$ and $\mu_2 + \nu_2$ are $\max\{K_1, K_2\}$-equivalent.

Proof. The implications (2) $\to$ (1) and (3) $\to$ (1) are elementary. We prove (1) $\to$ (2). This also proves (1) $\to$ (3) since one can take $X = (0, \mu_1(\mathbf{R})]$.

Note that (23) implies that the total μ_1 and μ_2 measures of $\mathbf{R}$ coincide, that is $\mu_1(\mathbf{R}) = \mu_2(\mathbf{R})$. Set $I = (0, \mu_1(\mathbf{R})]$. Define ψ_i as follows. Set

$$E_i(x) = \{y \in \mathbf{R} : \mu_i(-\infty, y] \geq x\},$$

for $i = 1, 2$. Let $\psi_i(x) = \inf E_i(x) = \min E_i(x)$. The fact the infimum of the set $E_i(x)$ belongs to this set follows from the countable additivity property of measures. In particular,

$$\int_{\infty}^{\psi_i(x)} d\mu_i \geq x.$$

Note that ψ_i is non-decreasing. It is elementary to check that $(\psi_i)_*(\nu) = \mu_i$. We now show that $||\psi_1 - \psi_2||_\infty \leq K$. Let $x \in I$. It follows from (23) that

$$x \leq \int_{-\infty}^{\psi_1(x)} d\mu_1 \leq \int_{-\infty}^{\psi_1(x)+K} d\mu_2.$$

This shows that $(\psi_1(x) + K) \in E_2(x)$. We conclude $\psi_2(x) \leq \psi_1(x) + K$. Similarly one shows $\psi_1(x) \leq \psi_2(x) + K$. This proves the proposition. □

For $\gamma \in \Gamma(S)$ let $\phi : \mathbf{R} \to \gamma$ be an isometric parametrisation (from the Euclidean metric on $\mathbf{R}$ to the hyperbolic on γ). Every other isometric parametrisation is obtained by pre-composing ϕ by a translation on $\mathbf{R}$. Let $\alpha \in \mathcal{M}(\mathbf{N}^1\Gamma(S))$ and let $\gamma^* \in \Gamma^*(S)$. The restriction of the measure α to $N^1\gamma^*$ is denoted by α_{γ^*}. Same as before, we identify (in the obvious way), the manifold $N^1\gamma^*$ and γ. Therefore, the pull-back measure $\phi^*\alpha_{\gamma^*}$ is well defined.

Definition 4.3. *Let $\alpha, \alpha' \in \mathbf{N}^1\Gamma(S)$ and let $K > 0$. For $\gamma \in \Gamma(S)$ let $\gamma^*(1), \gamma^*(2) \in \Gamma^*(S)$ denote the two orientations on γ. Let $\phi : \mathbf{R} \to \gamma$ be any isometric parametrisation. We have*

- *If for every $\gamma \in \Gamma(S)$ the measures $\phi^*\alpha_{\gamma^*(i)}$ and $\phi^*\alpha'_{\gamma^*(i)}$, $i = 1, 2$, are K-equivalent, then we say that the measures α and α' are K-equivalent.*

- *If for some $\gamma \in \Gamma(S)$ the measures $\phi^*\alpha_{\gamma^*(1)}$ and $\phi^*\alpha_{\gamma^*(2)}$ are K-equivalent, we say that the measure α is K-symmetric on γ. If α is K-symmetric on every $\gamma \in \Gamma(S)$ then we say that α is K-symmetric.*

The above definition does not depend on the choice of the parametrisation ϕ (see the above remark).

The following propositions are needed in the proof of Theorem 1.1 below.

Proposition 4.2. *Let $0 < \epsilon < 1$. Let $\alpha_1, \alpha_2 \in \mathcal{M}(\mathbf{R})$ be discrete measures with finitely many non-trivial atoms, and suppose that α_1 and α_2 are K-equivalent. Then there are measures $\alpha_i^{\text{rat}}, \alpha_i' \in \mathcal{M}(\mathbf{R})$, $i = 1, 2$, so that $\alpha_i^{\text{rat}} + \alpha_i' = \alpha_i$ and α_1^{rat} and α_2^{rat} are K equivalent. Also, α_i^{rat} has atoms of rational weights, and the weight of any atom of α_i' is at most ϵ.*

Proof. Let $a_i \in \mathbf{R}$ and $b_j \in \mathbf{R}$ denote respectively the points where α_1 and α_2 have non-trivial atoms. Set $x_i = \alpha_1(a_i)$ and $y_j = \alpha_2(b_j)$. Let m_1 be the total number of atoms a_i and let m_2 be the total number of atoms b_j. Set $m = m_1 + m_2$. Also, let A be the minimum of all non-zero weights of atoms of both measures, and B the maximum of all non-zero weights.

Since α_1 and α_2 are K-equivalent, we have that (23) holds. Since α_1 and α_2 have finitely many atoms, the condition (23) becomes a finite systems of linear inequalities with integer coefficients, in x_i and y_j. Each such inequality has the form

$$\sum_i \sigma_1(i) x_i \le \sum_j \sigma_2(j) y_j,$$

or

$$\sum_j \sigma_2(j) y_j \le \sum_i \sigma_1(i) x_i,$$

where every $\sigma_1(i)$, $\sigma_2(j)$ is either 1 or 0. If we treat x_i and y_j as real variables, then we conclude that this system of linear inequalities has a non-trivial solution. In fact, the set of solutions of each above inequality is a half-space in $\mathbf{R}^m$ (each half space contains the origin in $\mathbf{R}^m$). The set of the solutions of the entire system is the intersection of all these half-spaces in $\mathbf{R}^m$. We denote this set by Sol. Let $N \in \mathbf{N}$ and let $\text{Sol}_N = \text{Sol} \cap \{|x_i|, |y_j| \le N\}$. We have that Sol_N is a convex polyhedron (possibly degenerate) in $\mathbf{R}^m$. By the Krein-Milman theorem, Sol_N is the closure of the convex combinations of the extreme points on Sol_N (there are finitely many extreme points). Each extreme point is the unique solution of a certain system of equations with integer coefficients. Therefore, every extreme point is a rational point in $\mathbf{R}^m$. We conclude that the rational points in $\mathbf{R}^m$ are dense in every Sol_N and therefore the rational points are dense in $\text{Sol} = \cup_{N \in \mathbf{N}} \text{Sol}_N$.

If x_i and y_j is the only solution to this system, then these numbers have to be rational. If this is not the only solution, then we can choose rational solutions x_i^{rat} and y_j^{rat} to be as close to x_i and y_j as we want. Fix any $\epsilon_1 > 0$ and let x_i^{rat} and y_j^{rat} be rational numbers that satisfy all the above inequalities, and such that $A\epsilon_1 > |x_i - x_i^{\text{rat}}|$ and $A\epsilon_1 > |y_j - y_j^{\text{rat}}|$. Let t be a rational number so that

$$1 - 2\frac{\epsilon_1}{A} < t < 1 - \frac{\epsilon_1}{A}.$$

Then $tx_i^{\mathrm{rat}} < x_i$ and $ty_j^{\mathrm{rat}} < y_j$ and also tx_i^{rat} and ty_j^{rat} satisfy all the above inequalities. Moreover, the following inequalities hold

$$|x_i - x_i^{\mathrm{rat}}| < \epsilon_1\big(\frac{2B}{A} + A\epsilon_1 + 2{\epsilon_1}^2\big),$$

and

$$|y_j - y_j^{\mathrm{rat}}| < \epsilon_1\big(\frac{2B}{A} + A\epsilon_1 + 2{\epsilon_1}^2\big).$$

Choose ϵ_1 small enough so that $\epsilon_1\big(\frac{2B}{A} + A\epsilon_1 + 2{\epsilon_1}^2\big) < \epsilon$.

Let α_1^{rat} be the measure with the same non-trivial atoms as α_1 and $\alpha_1^{\mathrm{rat}}(a_i) = tx_i^{\mathrm{rat}}$. Similarly define $\alpha_2^{\mathrm{rat}}(b_j) = ty_j^{\mathrm{rat}}$. Set $\alpha_1' = \alpha_1 - \alpha_1^{\mathrm{rat}}$ and $\alpha_2' = \alpha_2 - \alpha_2^{\mathrm{rat}}$. The measures α_1' and α_2' are non-negative, and the weight of any atom under either of these two measures is less than ϵ. Since tx_i^{rat} and ty_j^{rat} satisfy the same inequalities as x_i and y_j we conclude that α_1^{rat} and α_2^{rat} are K-equivalent. □

Let L be a finite set. By $\vartheta_L \in \mathcal{M}(L)$ we denote the counting measure, that is if $L_1 \subset L$ then $\vartheta_L(L_1)$ is equal to the number of elements in L_1. On the other hand, we say that a Borel measure is integral if the measure of every set is an integer. The following proposition is left to the reader.

Proposition 4.3. *Let L be a finite set, and let $T : L \to A$ be a function. Suppose that $T_*(\vartheta_L) = \alpha_1 + \alpha_2$ such that $\alpha_1, \alpha_2 \in \mathcal{M}(A)$ are integral measures. Then we can write $L = L_1 \cup L_2$ where L_1 and L_2 are disjoint, such that $T_*(\vartheta_{L_1}) = \alpha_1$ and $T_*(\vartheta_{L_2}) = \alpha_2$.*

Proposition 4.4. *Let $\mathcal{L}_A$ and $\mathcal{L}_B$ be finite sets of labels, and let $\mathrm{lab}_A : \mathcal{L}_A \to \mathbf{R}$ and $\mathrm{lab}_B : \mathcal{L}_B \to \mathbf{R}$ be labelling maps. Suppose that the measures $\alpha = (\mathrm{lab}_A)_*(\vartheta_{\mathcal{L}_A})$ and $\beta = (\mathrm{lab}_B)_*(\vartheta_{\mathcal{L}_B})$ are K-equivalent. Then we can find a bijection $\sigma_{A,B} : \mathcal{L}_A \to \mathcal{L}_B$ such that $|| \mathrm{lab}_A - \mathrm{lab}_B \circ \sigma_{A,B}||_\infty \leq K$.*

Proof. Since α and β are K-equivalent, we conclude that the total mass of α is equal to the total mass of β. Set $m = \alpha(\mathbf{R}) = \beta(\mathbf{R})$. By Proposition 4.1, we can find non-decreasing and left continuous functions $\psi_A, \psi_B : (0, m] \to \mathbf{R}$ such that $||\psi_A - \psi_B||_\infty \leq K$ where $(\psi_l)_*(\nu) = (\mathrm{lab}_l)_*(\vartheta_{\mathcal{L}_l})$ for $l = A, B$ (here ν is the Lebesgue measure on $(0, m]$). Since $(\mathrm{lab}_l)_*(\vartheta_{\mathcal{L}_l})$ is an integral measure, we have that $\psi_l(t) = \psi_l(t^*)$ where t^* is the least integer greater or equal to t. Therefore, we can find bijections $\phi_l \to \{1, 2, ..., m\}$, $l = A, B$, such that $\psi_l(\phi_l(x)) = \mathrm{lab}_l(x)$ for every $x \in \mathcal{L}_l$.

For $a \in \mathcal{L}_A$ we let $\sigma_{A,B}(a) = (\phi_B)^{-1}(\phi_A(a))$. Then $| \mathrm{lab}_A(a) - \mathrm{lab}_B(\sigma_{A,B}(a))| = |\psi_A(\phi_A(a)) - \psi_B(\phi_A(a))| \leq K$ which proves the proposition. □

Proposition 4.5. *Let A and B be finite sets of real numbers. Let $\mathcal{L}_A$ and $\mathcal{L}_B$ be finite sets of labels, and $\mathrm{lab}_A : \mathcal{L}_A \to A$ and $\mathrm{lab}_B : \mathcal{L}_B \to B$ labelling maps. Let $\alpha, \beta \in \mathcal{M}(\mathbf{R})$ be the atomic measures, that are supported on A and B respectively, and $\alpha(a) = | \mathrm{lab}_A^{-1}(a)|$ for every $a \in A$ and $\beta(b) = | \mathrm{lab}_B^{-1}(b)|$ for every $b \in B$ (here $| \mathrm{lab}_A^{-1}(a)|$, $| \mathrm{lab}_B^{-1}(b)|$ denote the numbers of the corresponding preimages in $\mathcal{L}_A$ and $\mathcal{L}_B$ respectively). Suppose that $\alpha = \alpha_1 + \alpha_2$ and $\beta = \beta_1 + \beta_2$ where $\alpha_i, \beta_j \in \mathcal{M}(\mathbf{R})$ are integral measures, such that α_1, β_1 are K_1-equivalent, and α_2, β_2 are K_2-equivalent. Denote by m_1 the total mass of α_1 and by m_2 the total mass of α_2. Then, there is a bijection $\sigma_{A,B} : \mathcal{L}_A \to \mathcal{L}_B$ so that for m_1 elements $a \in \mathcal{L}_A$ we have $| \mathrm{lab}_B(\sigma_{A,B}(a)) - \mathrm{lab}_A(a)| < K_1$ and for the remaining m_2 elements $b \in \mathcal{L}_A$ we have $| \mathrm{lab}_B(\sigma_{A,B}(b)) - \mathrm{lab}_A(b)| < K_2$.*

Proof. Note that $\alpha = (\mathrm{lab}_A)_*(\vartheta_{\mathcal{L}_A})$ and $\beta = (\mathrm{lab}_B)_*(\vartheta_{\mathcal{L}_B})$. The proof follows from Proposition 4.3 and Proposition 4.4. □

We end this subsection with the following two elementary propositions that will be used in Section 6.

Proposition 4.6. *Let $\alpha, \beta, \eta \in \mathcal{M}(\mathbf{R})$ such that α and β are K-equivalent, and such that β and η are L-equivalent. Then α and η are $K + L$-equivalent.*

Proof. One directly verifies that the condition (1) from Proposition 4.1 holds for the measures α and η. □

Proposition 4.7. *Let (X, ν) be a measure space and let $\mu_i : X \to \mathcal{M}(\mathbf{R})$, $i = 1, 2$, be such that the mapping $x \to \mu_i(x)(-\infty, t]$ is measurable for every $t \in \mathbf{R}$. Suppose that $\mu_1(x)$ and $\mu_2(x)$ are K-equivalent for every $x \in X$. Then the measures*

$$\mu_1(X) = \int\limits_X \mu_1(x)\, d\nu(x),$$

and

$$\mu_2(X) = \int\limits_X \mu_2(x)\, d\nu(x),$$

are K-equivalent.

Remark. If $\mu : X \to \mathcal{M}(N^1\mathbf{R})$ is a measurable mapping, and if $\mu(x)$ is K-symmetric for every x, then the measure

$$\mu(X) = \int\limits_X \mu(x)\, d\nu(x),$$

is also K-symmetric.

Proof. One verifies directly that the condition (1) from Proposition 4.1 holds for the measures $\mu_1(X)$ and $\mu_2(X)$. □

4.3. **Proof of Theorem 1.1.** The remainder of the paper after this subsection is devoted to proving the next theorem. From now on $P(r)$ denotes a polynomial that only depends on S. In particular we have $P(r) + P(r) = P(r)$ and $rP(r) = P(r)$.

Theorem 4.1. *There exist a constant $r_0(S) = r_0$ that depends only on S, so that for every $r > r_0$ there exists a finite measure $\mu(r) \in \mathcal{M}(\mathcal{T}(S))$ so that the total measure $|\mu(r)|$ satisfies the inequality*

$$\frac{\Lambda(T^1S)}{2} < |\mu(r)| < \frac{3\Lambda(T^1S)}{2},$$

and with the following properties. There exist measures $\alpha(r), \alpha_1(r), \beta(r) \in \mathcal{M}(\mathbf{N}^1\Gamma(S))$ so that the measure $\widehat{\partial}\mu(r) \in \mathcal{M}(\mathbf{N}^1\Gamma(S))$ can be written as $\widehat{\partial}\mu(r) = \alpha(r) + \alpha_1(r) + \beta(r)$ and the following holds

(1) *Let $\widehat{\mu}(r)$ denote the restriction of the measure $\mu(r)$ to the set of triangles $T \in \mathcal{T}(S)$ for which $\mathbf{h}(T) \geq r^2$. Then*

$$\int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1) d\widehat{\partial}\widehat{\mu}(r) \leq P(r)e^{-r}.$$

(2) *The measure* $\alpha(r)$ *is* re^{-r}*-symmetric, and the measure* $\alpha_1(r)$ *is* Q*-symmetric, for any* $Q > 200$.
(3) *We have*

$$\int_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1) d\beta(r) \leq P(r)e^{-r}.$$

(4) *We have*

$$\int_{\mathbf{N}^1\Gamma(S)} d\alpha_1(r) \leq e^{-r}.$$

We will explicitly construct the required measure $\mu(r)$. In the remainder of this section we prove Theorem 1.1 assuming Theorem 4.1 and its notation. First, we show that we may assume that the measure $\mu(r)$ from Theorem 4.1, has finite support.

Proposition 4.8. *Assume that Theorem 4.1 holds. Then there exist a constant* $r_0(S) = r_0$ *that depends only on* S *so that for every* $r > r_0$ *there exists a finite measure* $\widetilde{\mu}(r) \in \mathcal{M}(\mathcal{T}(S))$ *with finite support, so that the total measure* $|\widetilde{\mu}(r)|$ *satisfies the inequality*

$$\frac{\Lambda(T^1S)}{4} < |\widetilde{\mu}(r)| < 2\Lambda(T^1S),$$

and with the following properties. There exist measures $\widetilde{\alpha}(r), \widetilde{\alpha}_1(r), \widetilde{\beta}(r) \in \mathcal{M}(\mathbf{N}^1\Gamma(S))$ *so that the measure* $\widehat{\partial}\widetilde{\mu}(r) \in \mathcal{M}(\mathbf{N}^1\Gamma(S))$ *can be written as* $\widehat{\partial}\widetilde{\mu}(r) = \widetilde{\alpha}(r) + \widetilde{\alpha}_1(r) + \widetilde{\beta}(r)$ *and the following holds*

(1) *The measure* $\widetilde{\alpha}(r)$ *is* re^{-r}*-symmetric, and the measure* $\widetilde{\alpha}_1(r)$ *is* Q*-symmetric, for any* $Q > 200$.
(2) *We have*

$$\int_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1) d\widetilde{\beta}(r) \leq P(r)e^{-r}.$$

(3) *We have*

$$\int_{\mathbf{N}^1\Gamma(S)} d\widetilde{\alpha}_1(r) \leq e^{-r}.$$

Proof. Let $\widetilde{\mu}(r) \in \mathcal{M}(\mathcal{T}(S))$ be such that $\mu(r) = \widetilde{\mu}(r) + \widehat{\mu}(r)$ where $\widehat{\mu}(r)$ was defined in the statement of Theorem 4.1. Then $\widetilde{\mu}(r)$ has finite support since it is supported only on triangles from $\mathcal{T}(S)$ which satisfy $\mathbf{h}(T) \leq r^2$ (and there are only finitely many such triangles). By the assumption in Theorem 4.1 we have $|\widehat{\mu}(r)| \leq P(r)e^{-r}$. This shows that

$$\frac{\Lambda(T^1S)}{4} < |\widetilde{\mu}(r)| < 2\Lambda(T^1S),$$

for r large enough. Note that the measure $\widehat{\partial}\mu(r)$ consists of countably many atoms, that are obtained as the feet of the centres of the triangles from $\mathcal{T}(S)$. The measure $\widetilde{\mu}(r)$ consists of finitely many atoms. It remains to construct the decomposition $\widehat{\partial}\widetilde{\mu}(r) = \widetilde{\alpha}(r) + \widetilde{\alpha}_1(r) + \widetilde{\beta}(r)$

For a discrete positive measure η we let $\operatorname{supp}(\eta)$ be the set of points $t \in \mathbf{R}$ for which $\eta(\{t\}) > 0$. We observe that $\widehat{\mu}$ and $\widetilde{\mu}$ have disjoint support because if T_1, T_2

are triangles that share a geodesic γ, and such that $\text{foot}_\gamma(\text{ct}(T_1)) = \text{foot}_\gamma(\text{ct}(T_2))$ then $T_1 = T_2$ (recall that foot_γ is an element of $N^1\gamma$ so if $\text{foot}_\gamma(\text{ct}(T_1)) = \text{foot}_\gamma(\text{ct}(T_2))$ then T_1 and T_2 are on the same side of γ). We let $\underline{\alpha}$ be the restriction of α to $\text{supp}\widetilde{\mu}$ and $\widehat{\alpha}$ the restriction of α to $\text{supp}\widehat{\mu}$. One defines $\underline{\alpha}_1$, $\widehat{\alpha}_1$, $\underline{\beta}$ and $\widehat{\beta}$ likewise. Then

$$\widetilde{\mu} = \underline{\alpha} + \underline{\alpha}_1 + \underline{\beta}, \qquad \widehat{\mu} = \widehat{\alpha} + \widehat{\alpha}_1 + \widehat{\beta},$$

and $\underline{\alpha} + \widehat{\alpha} = \alpha$, $\underline{\alpha}_1 + \widehat{\alpha}_1 = \alpha_1$ and $\underline{\beta} + \widehat{\beta} = \beta$.

We aim to "symmetrise" $\underline{\alpha}$ so that it is re^{-r}-symmetric. To this end for each $\gamma \in \Gamma(S)$ we choose an orientation $\gamma^* \in \Gamma^*(S)$ for γ. Let α^+ be the restriction of α to $N^1\gamma^*$ and α^- be the restriction of α to $N^1(-\gamma^*)$, and think of α^+ and α^- as measures on γ. We define $\underline{\alpha}^+$, $\underline{\alpha}^-$, $\widehat{\alpha}^+$ and $\widehat{\alpha}^-$ likewise. Since α is re^{-r}-symmetric by Proposition 4.1 we can write

$$\alpha^{+/-} = (\psi_{+/-})_*[0, \alpha^+(\gamma)),$$

such that $\psi_{+/-} : [0, \alpha^+(\gamma)) \to \gamma$ satisfy that $\mathbf{d}(\psi_+(t), \psi_-(t)) \le re^{-r}$ for all $t \in [0, \alpha^+(\gamma))$. Then we let $A_{+/-} \subset [0, \alpha^+(\gamma))$ be defined by $A_{+/-} = \psi^{-1}_{+/-}(\text{supp}(\alpha_{+/-}))$. Set $A = A_+ \cap A_-$. Also, let $\widetilde{\alpha}^{+/-} = (\psi_{+/-})_* A$. By construction $\widetilde{\alpha}^+$ and $\widetilde{\alpha}^-$ are re^{-r}-equivalent. Now consider $\widetilde{\alpha}^+$ and $\widetilde{\alpha}^-$ as measures on $N^1\gamma^*$ and $N^1(-\gamma^*)$ respectively. Let $\widetilde{\alpha} = \widetilde{\alpha}^+ + \widetilde{\alpha}^-$. Then $\widetilde{\alpha}$ is re^{-r}-symmetric.

We let $\eta = \underline{\alpha} - \widetilde{\alpha}$. The reader can verify that there exists a measure $\eta' \le \widehat{\alpha}$ such that $\eta' + \eta$ is re^{-r}-symmetric (the measure η' is constructed in the obvious way using the maps $\psi_{+/-}$). Note that if $(\gamma^*, z), (\gamma^*, z') \in N^1\gamma$ satisfy $\mathbf{d}(z, z') \le K$ then $\mathcal{K}(\gamma^*, z) \le e^K \mathcal{K}(-\gamma^*, z_1)$. It follows that

$$\int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1)\, d\eta \le e^{re^{-r}} \int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1)\, d\eta' \le$$

$$\le e^{re^{-r}} \int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1)\, d\widehat{\alpha} \le 3 \int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1)\, d\widehat{\alpha},$$

for r large enough.

We likewise find $\widetilde{\alpha}_1 \le \underline{\alpha}_1$ such that $\widetilde{\alpha}_1$ is Q-symmetric and $\eta_1 = \underline{\alpha}_1 - \widetilde{\alpha}_1$ satisfy

$$\int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1)\, d\eta_1 \le e^Q \int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1)\, d\widehat{\alpha}_1.$$

Also

$$\int\limits_{\mathbf{N}^1\Gamma(S)} d\widetilde{\alpha}_1 \le e^{-r}.$$

This shows that $\widetilde{\alpha}$ and $\widetilde{\alpha}_1$ satisfy the conditions (1) and (3) of this proposition.

We then let $\widetilde{\beta} = \underline{\beta} + \eta + \eta_1$. Then $\widetilde{\mu} = \widetilde{\alpha} + \widetilde{\alpha}_1 + \widetilde{\beta}$. Moreover

$$\int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z))\, d\widetilde{\beta} = \int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z))\, d(\underline{\beta} + \eta + \eta_1) \le$$

$$\le \int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z))\, d\beta + \le \int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z))\, d(\widehat{\alpha} + \widehat{\alpha}_1) \le P(r)e^{-r},$$

since $\widehat{\alpha}+\widehat{\alpha}_1 \leq \widehat{\mu}$. This proves the proposition.

□

The idea is to use the measures $\widetilde{\mu}(r)$ to construct certain admissible pairs $(\mathcal{C}(r), \sigma_{\mathcal{C}(r)})$. In turn, this will enable us to construct the corresponding covers of S that are required by Theorem 1.1.

From now on we suppress the dependence on r that is we set $\widetilde{\mu}(r)=\mu$, $\widetilde{\alpha}(r)=\alpha$, $\widetilde{\alpha}_1(r)=\alpha_1$ and $\widetilde{\beta}(r)=\beta$. Since μ is finitely supported, the measure μ is atomic and it has finitely many atoms. The measure $\widehat{\partial}\mu$ is atomic and it has finitely many atoms, as well. Let m denote the total number of non-trivial atoms for $\widehat{\partial}\mu$. Let L be the maximum of the combinatorial length function $\mathcal{K}$ over the m points where $\widehat{\partial}\mu$ is supported. Let

$$\epsilon=\frac{1}{m(L+1)e^r}.$$

For $T \in \mathcal{T}(S)$ choose a number $0 \leq \mu'(T) \leq \epsilon$ so that $\mu'(T)+\mu(T)$ is a rational number, and set $\mu'(T)+\mu(T)=\mu^{\text{rat}}(T)$. If $\mu(T)=0$ then we set $\mu'(T)=0$ as well. This is how we define two new measures $\mu', \mu^{\text{rat}} \in \mathcal{M}(\mathcal{T}(S))$. Both these measures are atomic (finitely many atoms), and μ^{rat} has the same set of atoms as μ (the weights of atoms of μ^{rat} are rational numbers). In particular, μ' has at most m atoms, and each of them has the weight at most ϵ. Note that the total measure of μ^{rat} satisfies $\mu(\mathcal{T}(S)) \geq \mu^{\text{rat}}(\mathcal{T}(S)) \geq \mu(\mathcal{T}(S))-m\epsilon > \mu(\mathcal{T}(S))-e^{-r}$

The measure $\widehat{\partial}\mu \in \mathcal{M}(\mathbf{N}^1\Gamma(S))$ is also atomic, with finitely many atoms as well. Let $\gamma \in \Gamma(S)$ and let $\gamma^*_1, \gamma^*_2 \in \Gamma^*(G)$ denote the two orientations on γ. Since α is re^{-r}-symmetric, the two measures on $\mathbf{R}$ that arise as the restrictions of α on $N^1\gamma^*_1$ and $N^1\gamma^*_2$ respectively, are atomic (finitely many atoms), re^{-r}-equivalent measures. Applying Proposition 4.2 on these two measures, and repeating the same for each $\gamma \in \Gamma(S)$ we construct the measures $\alpha^{\text{rat}}, \alpha' \in \mathcal{M}(N^1\Gamma(S))$ with the properties:

- $\alpha^{\text{rat}}+\alpha'=\alpha$.
- α^{rat} is re^{-r}-symmetric.
- for any atom $V \in \mathbf{N}^1\Gamma(S)$ of α' we have $\alpha'(V) \leq \epsilon$.

We repeat the same for α_1. We construct the measures $\alpha^{\text{rat}}_1, \alpha'_1 \in \mathcal{M}(N^1\Gamma(S))$ with

- $\alpha^{\text{rat}}_1+\alpha'_1=\alpha_1$.
- α^{rat} is Q-symmetric.
- for any atom $V \in \mathbf{N}^1\Gamma(S)$ of α'_1 we have $\alpha'_1(V) \leq \epsilon$.

We have

$$\widehat{\partial}\mu^{\text{rat}}=\widehat{\partial}\mu+\widehat{\partial}\mu'=\alpha^{\text{rat}}+{\alpha_1}^{\text{rat}}+(\beta+\widehat{\partial}\mu'+\alpha'+\alpha'_1).$$

Note that from the fourth inequality in Proposition 4.6, we have

$$\int\limits_{\mathbf{N}^1\Gamma(S)} d\alpha^{\text{rat}}_1 \leq \int\limits_{\mathbf{N}^1\Gamma(S)} d\alpha_1 \leq e^{-r}.$$

Set $\beta_1=\beta+\widehat{\partial}\mu'+\alpha'+\alpha'_1$. Since μ^{rat} is atomic (finitely many atoms) with rational weights, so is the measure $\widehat{\partial}\mu^{\text{rat}}$. Since α^{rat} and ${\alpha_1}^{\text{rat}}$ are also atomic (finitely many atoms) with rational weights, and since all the measure in question are positive, we conclude that β_1 is also atomic (finitely many atoms) and with rational weights. Moreover, we have

$$\int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1) d\beta_1 \leq \int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1) d\beta + \int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1) d\widehat{\partial}\mu' +$$

$$+ \int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1) d\alpha' + \int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1) d\alpha_1' \leq P(r)e^{-r} + 3m(L+1)\epsilon,$$

that is, for r large enough we have

$$\int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1) d\beta_1 \leq P(r)e^{-r}.$$

Let n be a large enough integer so that the weights of atoms of the measures μ^{rat}, α^{rat}, ${\alpha_1}^{\mathrm{rat}}$ and β_1 are all integers. We multiply the measures in question by n. Set $\mu^{\mathrm{int}} = n\mu^{\mathrm{rat}}$, $\alpha^{\mathrm{int}} = n\alpha^{\mathrm{rat}}$, $\alpha_1^{\mathrm{int}} = n{\alpha_1}^{\mathrm{rat}}$ and $\beta_1^{\mathrm{int}} = n\beta_1$. We have that α^{int} is still re^{-r}-symmetric, and α_1^{int} is Q-symmetric. Moreover, we have the following estimate for the total measure of α_1^{int}

$$\int\limits_{\mathbf{N}^1\Gamma(S)} d\alpha_1^{\mathrm{int}} \leq ne^{-r}. \tag{24}$$

Also, the total measure of μ^{int} satisfies that $n\mu(\mathcal{T}(S)) \geq \mu^{\mathrm{int}}(\mathcal{T}(S)) > n(\mu(\mathcal{T}(S)) - e^{-r})$.

Now we apply the Correction lemma (Lemma 3.3). Let $(\gamma^*, z) \in \mathbf{N}^1\Gamma(S)$ be an atom of β_1^{int} (note that the same point can be an atom of the measure α^{int} or α_1^{int}). Choose a lift of (γ^*, z) to $\mathbf{N}^1\Gamma(G)$ and denote it also by (γ^*, z). We apply Lemma 3.3 to this pair, to obtain the corresponding polygon $\mathcal{P}$ in $\mathbf{H}$ that is to the right of γ^*. Project the triangles from the triangulation $\tau(\mathcal{P})$ to S. Let $T' \in \tau(\mathcal{P})$ and let $[T']_G = T \in \mathcal{T}(S)$ be its projection. Let $\nu^T \in \mathcal{M}(\mathcal{T}(S))$ be the measure that is supported on T and so that $\nu^T(T) = \beta_1^{\mathrm{int}}(\gamma^*, z)$. Set

$$\nu^{(\gamma^*, z)} = \sum_T \nu^T,$$

where we sum over all such triangles T. From Lemma 3.3 we have the bound on the number of triangles in $\tau(\mathcal{P})$ and thus we obtain the following estimate of the total measure

$$\nu^{(\gamma^*, z)}(\mathcal{T}(S)) \leq \big(C\mathcal{K}(\gamma^*, z) + K\big)\beta_1^{\mathrm{int}}(\gamma^*, z).$$

Here C and K are the constants from Lemma 3.3. Note that each atom of $\nu^{(\gamma^*, z)}$ has an integer weight.

Let D be the constant from Lemma 3.3, and let $\gamma_0 \in \Gamma(S) \setminus \lambda_{\mathrm{Gen}}(S)$ be a geodesic that lifts to an edge of a triangle from $\tau(\mathcal{P})$. Denote by $\beta_1^{\mathrm{int}}|_{(\gamma^*, z)}$ the restriction of β_1^{int} to the point $(\gamma^*, z) \in N^1\Gamma(S)$. Then the two measures on γ_0 that are the restrictions of the measure $\widehat{\partial}\nu^{(\gamma^*, z)} + \beta_1^{\mathrm{int}}|_{(\gamma^*, z)}$ on $N^1\gamma_0$ are D-equivalent. This follows from Lemma 3.3, that is, the two atoms of $\widehat{\partial}\nu^{(\gamma^*, z)} + \beta_1^{\mathrm{int}}|_{(\gamma^*, z)}$ in $N^1\gamma_0$ (one on each side of γ_0), are within the hyperbolic distance D (these two atoms have the same weight by the definition of $\nu^{(\gamma^*, z)}$).

Repeat this process for every non-trivial atom of β_1^{int} and set

$$\nu = \sum_{(\gamma^*,z)} \nu^{(\gamma^*,z)},$$

where we sum over all non-trivial atoms for β_1^{int}. We have $\nu \in \mathcal{M}(\mathcal{T}(S))$ and

$$\nu(\mathcal{T}(S)) \leq \int_{\mathbf{N}^1\Gamma(S)} \big(C\mathcal{K}(\gamma^*,z)+K\big) d\beta_1^{\text{int}} \leq nP(r)e^{-r}.$$

Let $\gamma_0 \in \Gamma(S) \setminus \lambda_{\text{Gen}}(S)$. Then the two measures on γ_0 that are the restrictions of the measure $\widehat{\partial}\nu + \beta_1^{\text{int}}$ to $N^1\gamma_0$ are D-equivalent. Let $\gamma_0 \in \lambda_{\text{Gen}}(S)$. Then by Lemma 3.3 all the atoms of $\widehat{\partial}\nu$ on γ_0 are within the $D/2$ hyperbolic distance from the point $\text{mid}(\gamma_0) \in \gamma_0$. If we can show that the total measures are equal, that is of $\widehat{\partial}\nu(N^1\gamma_0^*) = \widehat{\partial}\nu(N^1(-\gamma_0^*))$ that would show that the measure $\widehat{\partial}\nu + \beta_1^{\text{int}}$ is D-symmetric.

Set $\mu_1^{\text{int}} = \mu^{\text{int}} + \nu$. Again, μ_1^{int} has finitely many atoms, and all the weights are integers. The above estimate for $\nu(\mathcal{T}(S))$ implies that, for r large enough, the total measure of μ_1^{int} satisfies the following inequalities

$$\frac{n\mu(\mathcal{T}(S))}{2} < n(\mu(\mathcal{T}(S)) - e^{-r}) < \mu_1^{\text{int}}(\mathcal{T}(S)) < n\mu(\mathcal{T}(S)) + n\mu(\mathcal{T}(S))P(r)e^{-r} < 2n\mu(\mathcal{T}(S)). \tag{25}$$

Set $\beta_2^{\text{int}} = \beta_1^{\text{int}} + \widehat{\partial}\nu$. Then the following equality holds

$$\widehat{\partial}\mu_1^{\text{int}} = \alpha^{\text{int}} + \alpha_1^{\text{int}} + \beta_2^{\text{int}}.$$

We have the following estimates on the total measures of β_2^{int}

$$\beta_2^{\text{int}}(\mathcal{T}(S)) \leq \int_{\mathbf{N}^1\Gamma(S)} d\nu + \int_{\mathbf{N}^1\Gamma(S)} d\beta_1^{\text{int}} < nP(r)e^{-r}. \tag{26}$$

Since μ_1^{int} has finitely many atoms with integer weights, we can consider μ_1^{int} as an element of $\mathbf{N}\mathcal{T}(S)$. We construct the labelled collection of triangles $\mathcal{C}$ that corresponds to $\mu_1^{\text{int}} \in \mathbf{N}\mathcal{T}(S)$ (see remark before Definition 3.1) . Let M denote the total number of elements of $\mathcal{L}_{\mathcal{C}}$ (clearly M is equal to the total measure of $3\mu_1^{\text{int}}$). From (25) we have $\frac{3n\mu(\mathcal{T}(S))}{2} < M < 6n\mu(\mathcal{T}(S))$.

Fix $\gamma \in \Gamma(S) \setminus \lambda_{\text{Gen}}(S)$ and choose an orientation γ^* on γ. We have already seen that the restriction of the measure μ_1^{int} on $N^1\gamma$ can be written as the sum of three measures from $\mathcal{M}(N^1\gamma)$ where each of these three measures produces a pair of measures on γ that are equivalent (for some constant). This implies that the sets $\mathcal{L}_{\mathcal{C},\gamma^*}, \mathcal{L}_{\mathcal{C},-\gamma^*} \subset \mathcal{L}_{\mathcal{C}}$ have the same number of elements. This shows that such γ^* does not figure in the formal sum $\partial\mathcal{C} \in \mathbf{Z}\Gamma^*(S)$. So

$$\partial\mathcal{C} = \sum_{i=1}^{l} k_i \gamma_{k_i}^*,$$

where $\gamma_{k_i}^* \in \lambda_{\text{Gen}}^*(S)$ and $k_i \in \mathbf{Z}$. By Proposition 3.2 we have that every $k_i = 0$ that is $\partial\mathcal{C} = 0$ in $\mathbf{Z}\Gamma^*(S)$. Thus for every $\gamma \in \Gamma(S)$ the total measures of $\widehat{\partial}\mu_1^{\text{int}}$ on $N^1\gamma^*$ and on $N^1(-\gamma^*)$ are the same. As indicated above, this proves that the measure $\beta_2^{\text{int}} = \beta_1^{\text{int}} + \widehat{\partial}\nu$ is D-symmetric.

Set $\alpha_2^{\mathrm{int}} = \alpha_1^{\mathrm{int}} + \beta_2^{\mathrm{int}}$. Then α_2^{int} is $\max\{Q,D\}$-symmetric, and from (24), (26), for r large enough, we have

$$\alpha_2^{\mathrm{int}}(\mathcal{T}(S)) \le nP(r)e^{-r} + nP(r)e^{-r} \le nP(r)e^{-r}. \tag{27}$$

Let $\gamma \in \Gamma(S)$ and consider the restriction of the measure $\widehat{\partial}\mu_1^{\mathrm{int}}$ to $N^1\gamma$. This produces a pair of measures in $\mathcal{M}(\mathbf{R})$. Apply Proposition 4.5 to this pair of measures. Choose an orientation γ^* on γ. Recall the notation from Proposition 4.5. We say that $z \in A \subset \gamma$ if (γ^*, z) is a non-trivial atom of the measure $\widehat{\partial}\mu_1^{\mathrm{int}}$. We say that $z \in B \subset \gamma$ if $(-\gamma^*, z)$ is a non-trivial atom of the measure $\widehat{\partial}\mu_1^{\mathrm{int}}$. We identify $\mathcal{L}_A$ with $\mathcal{L}_{\mathcal{C},\gamma^*}$ and $\mathcal{L}_B$ with $\mathcal{L}_{\mathcal{C},-\gamma^*}$. We define the involution $\sigma_{\mathcal{C}} : \mathcal{L}_{\mathcal{C},\gamma^*} \to \mathcal{L}_{\mathcal{C},-\gamma^*}$ by $\sigma_{\mathcal{C}} = \sigma_{A,B}$ where $\sigma_{A,B} : A \to B$ is the bijection from Proposition 4.5. This is how we construct the admissible pair $(\mathcal{C}, \sigma_{\mathcal{C}})$.

Fix $a \in \mathcal{L}_{\mathcal{C}}$. Let $(T_1,\gamma_1^*), (T_1', -\gamma_1^*) \in \mathcal{T}^*(G)$ such that $\mathrm{lab}_{\mathcal{C}}(a) = [(T_1,\gamma_1^*)]$ and $\mathrm{lab}_{\mathcal{C}}(\sigma_{\mathcal{C}}(a)) = [(T_1', -\gamma_1^*)]$. Let $[T_1]_G = \mathrm{Pr}_S(\mathrm{lab}_{\mathcal{C}}(a)) = T \in \mathcal{T}(S)$ and $[T_1']_G = \mathrm{Pr}_S(\mathrm{lab}_{\mathcal{C}}(\sigma_{\mathcal{C}}(a))) = T' \in \mathcal{T}(S)$. Also, let $\gamma^* \in \Gamma^*(S)$ be the projection of γ_1^* to S and finally let $\gamma \in \Gamma(S)$ be the corresponding unoriented geodesic. We have $\mathbf{d}(\mathrm{foot}_\gamma(\mathrm{ct}(T)), \mathrm{foot}_\gamma(\mathrm{ct}(T'))) \le \max\{Q,D\}$ where $\mathrm{lab}_{\mathcal{C}}(\sigma_{\mathcal{C}}(a)) = T'$. Let $N_{\mathcal{C}}(re^{-r})$ denote the number of elements $a \in \mathcal{L}_{\mathcal{C}}$ so that for the corresponding edge γ of $T = \mathrm{Pr}_S(\mathrm{lab}_{\mathcal{C}}(a))$ we have $\mathbf{d}(\mathrm{foot}_\gamma(\mathrm{ct}(T)), \mathrm{foot}_\gamma(\mathrm{ct}(T'))) > re^{-r}$. From Proposition 4.5 and from (27), we conclude that $N_{\mathcal{C}}(re^{-r}) \le nP(r)e^{-r}$.

By Lemma 3.1, there exists finitely many virtual triangulation pairs $(\mathcal{C}_i, \sigma_i)$ so that $(\mathcal{C}, \sigma)$ is their union. For each i there exists a finite cover S_i of S so that $\mathcal{L}_{\mathcal{C}_i} = \tau^*(S_i)$ is a triangulation of S_i (by $\lambda(S_i)$ we denote the corresponding set of edges). Also, by $\mathbf{r}(i) \in X(\tau(S_i))$ we denote the corresponding shear coordinates. Note that the Riemann surface S_i corresponds to the point $F_{\tau(S_i)}(\mathbf{r}(i))$ in the corresponding Teichmüller space. For each S_i we have the following:

- Since $\mathbf{h}(T) \le r^2$ for every $T \in \tau(S_i)$ from Lemma 3.2 we have $O_{\tau(S_i)}(\mathbf{r}(i)) \le 2r^2$.
- We have $||\mathbf{r}(i)||_\infty < \max\{Q,D\}$.

One can verify the following elementary proposition.

Proposition 4.9. *Let $x_i, y_i > 0$, $i = 1,...,k$, and set $x = x_1 + ... + x_k$, $y = y_1 + ... + y_k$. Then*

$$\min_{1\le i\le k} \frac{x_i}{y_i} \le \frac{x}{y}.$$

Let $N_{\tau(S_i)}(re^{-r})$ denote the number of edges from $\lambda(S_i)$ for which the corresponding shear coordinate is greater than re^{-r}. We have

$$\sum_i N_{\tau(S_i)}(re^{-r}) = N_{\mathcal{C}}(re^{-r}) \le nP(r)e^{-r}.$$

Also,

$$\sum_i |\lambda(S_i)| = M \ge \frac{3n\mu(\mathcal{T}(S))}{2}.$$

From the above proposition we have that for at least one surface S_i, say for S_1, we have that

$$\frac{N_{\tau(S_1)}(re^{-r})}{|\lambda(S_1)|} \le P(r)e^{-r}.$$

where $|\lambda(S_1)|$ is the total number of edges in $\lambda(S_1)$.

Let $r \to \infty$. From Theorem 2.2 we have that the Weil-Petersson distance between $F_{\tau(S_i)}(\mathbf{r}(i))$ and $F_{\tau(S_i)}(0)$, tends to 0 when $r \to \infty$. Note that by Proposition 2.2, the Riemann surface that corresponds to $F_{\tau(S_i)}(0)$ is isomorphic to the quotient of $\mathbf{H}$ by a finite index subgroup of $\mathbf{PSL}(2,\mathbf{Z})$.

Theorem 1.1 states that for any two punctured Riemann surfaces S and R of finite type, and for every $\epsilon > 0$ we can find finite covers S_ϵ and R_ϵ of S and R respectively, so that the Weil-Petersson distance between them is less than ϵ. We first find S'_ϵ and R'_ϵ finite covers of S and R respectively, so that S'_ϵ and R'_ϵ are $\frac{\epsilon}{2}$-close (in the Weil-Petersson sense) to two Riemann surfaces S''_ϵ and R''_ϵ where S''_ϵ and R''_ϵ are isomorphic to $\mathbf{H}/G_1$ and $\mathbf{H}/G_2$ respectively, where G_1, G_2 are finite index subgroups of $\mathbf{PSL}(2,\mathbf{Z})$. Set $G_3 = G_1 \cap G_2$ and let M_ϵ be the corresponding Riemann surface. Then there are covers S_ϵ and R_ϵ of S and R respectively, so that S_ϵ and R_ϵ are $\frac{\epsilon}{2}$-close (in the Weil-Petersson sense) to M. Moreover, the Weil-Petersson distance between S_ϵ and R_ϵ is at most ϵ. This proves the theorem.

5. The absorption maps and the proof of Theorem 4.1

5.1. **Preliminary results from hyperbolic geometry and four operations on the unit tangent bundle.** Let $\mathbf{Mob}(\mathbf{H})$ denote the group of orientation preserving Möbius transformations of $\mathbf{H}$. By $\mathbf{Mob}_\infty(\mathbf{H})$ we denote the subgroup of $\mathbf{Mob}(\mathbf{H})$ whose elements preserve ∞. By $\Gamma(\mathbf{H})$ we denote the set of all geodesics in $\mathbf{H}$. Also, let $\mathcal{T}(\mathbf{H})$ denote the set of all ideal triangles in $\mathbf{H}$.

Denote by $T^1\mathbf{H}$ the unit tangent bundle of $\mathbf{H}$. Elements of $T^1\mathbf{H}$ are pairs (z,v) where $z \in \mathbf{H}$ and v is a unit vector at z. If $f \in \mathbf{Mob}(\mathbf{H})$ then $f(z,v) \in T^1\mathbf{H}$ is a well defined element. The quotient $T^1\mathbf{H}/G$ is isomorphic to the unit tangent bundle T^1S. Also, if $E \subset \mathbf{H}$ or $E \subset S$ then T^1E denotes the restriction of the corresponding unit tangent bundle over E.

We parametrise $T^1\mathbf{H}$ by $(z,v) = (x,y,\theta)$. Here $z = x+iy$, and $\theta \in [0,2\pi)$ is the positively oriented angle that v makes with the positive part of the x-axis. The Liouville volume form on $T^1\mathbf{H}$ is given by $d\Lambda = y^{-2}\, dx \wedge dy \wedge d\theta$. The corresponding measure on $T^1\mathbf{H}$ is called the Liouville measure. The total measure $\Lambda(T^1S)$ is finite.

Let $\omega = e^{2\pi i/3}$. By $\omega : T^1\mathbf{H} \to T^1\mathbf{H}$ we also denote the map $\omega(z,v) = (z,\omega v)$. Clearly the map ω is a diffeomorphism of $T^1\mathbf{H}$ of order three (this is the first operation on $T^1\mathbf{H}$ we define in this subsection). Also, the map ω commutes with every element from $\mathbf{Mob}(\mathbf{H})$ (and in particular ω commutes with every element of the group G). We have that $\omega : T^1S \to T^1S$ is a well defined diffeomorphism. Moreover, ω is measure preserving (it preserves the Liouville measure on $T^1\mathbf{H}$), that is $\mathrm{Jac}(\omega) = 1$ where $\mathrm{Jac}(\omega)$ denotes the Jacobian.

Fix $(z,v) \in T^1\mathbf{H}$. Let $\gamma_{(z,v)} : [0,\infty) \to \mathbf{H}$ be the natural parametrisation of the geodesic ray that starts at z and that is tangent to the vector v at z that is $\gamma'_{(z,v)}(0) = v$. We use the same notation for the induced map $\gamma_{(z,v)} : \mathbf{R}^+ \cup \{0\} \to S$. By $\gamma_{(z,v)}$ we also denote the corresponding geodesic ray $\gamma_{(z,v)}(\mathbf{R}^+ \cup \{0\})$.

As usual, for $t \in \mathbf{R}^+ \cup \{0\}$ the geodesic flow $\mathbf{g}_t : T^1\mathbf{H} \to T^1\mathbf{H}$ is given by $\mathbf{g}_t(z,v) = (\gamma_{(z,v)}(t), \gamma'_{(z,v)}(t))$. Moreover, for $f \in \mathbf{Mob}(\mathbf{H})$ we have $\mathbf{g}_t \circ f = f \circ \mathbf{g}_t$. Therefore, the flow is well defined on $T^1\mathbf{H}/G$. We use the same notation for the induced flow $\mathbf{g}_t : T^1S \to T^1S$. The geodesic flow $\mathbf{g}_t$ is measure preserving.

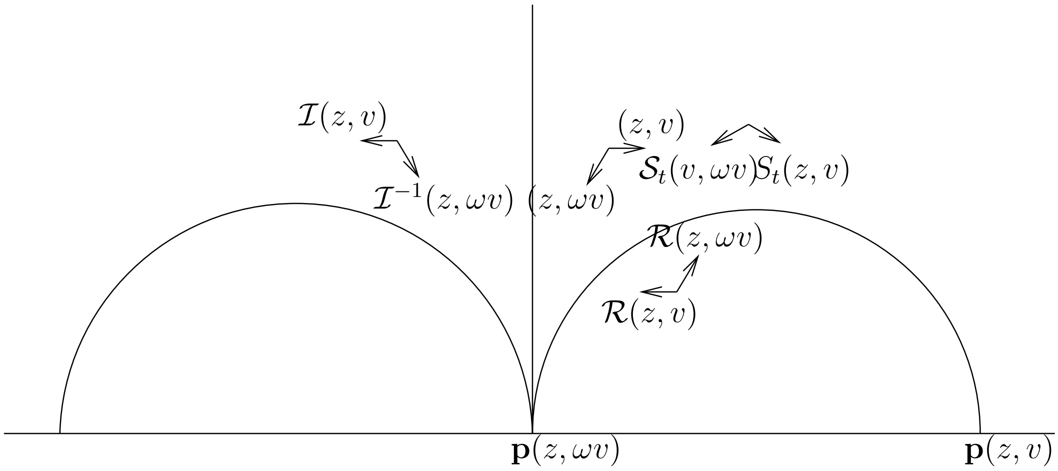

FIGURE 4. The four operations on the $T^1\mathbf{H}$

Definition 5.1. *Let* $\mathbf{p}: T^1\mathbf{H} \to \partial\mathbf{H}$ *denote the map such that* $\mathbf{p}(z,v) \in \partial\mathbf{H}$ *is the end point of the geodesic ray* $\gamma_{(z,v)}$. *Let* $\mathbf{p}^1 : T^1\mathbf{H} \to \Gamma(\mathbf{H})$ *denote the map such that* $\mathbf{p}^1(z,v) \in \Gamma(\mathbf{H})$ *is the geodesic that connects the points* $\mathbf{p}(z,v)$ *and* $\mathbf{p}(z,\omega v)$. *Let* $\mathbf{p}^2 : T^1\mathbf{H} \to \mathcal{T}(\mathbf{H})$ *denote the map such that* $\mathbf{p}^2(z,v) \in \mathcal{T}(\mathbf{H})$ *is the triangle with the vertices* $\mathbf{p}(z,v)$ $\mathbf{p}(\omega(\mathbf{z},v))$ *and* $\mathbf{p}(\omega^2(\mathbf{z},v))$.

Since the angle between v and ωv is $2\pi/3$ we find that the point z is the centre of the triangle $\mathbf{p}^2(z,v) = \mathbf{p}^2(\omega(z,v)) = \mathbf{p}^2(\omega^2(z,v))$. The triangle $\mathbf{p}^2(z,v)$ is bounded by the geodesics $\mathbf{p}^1(z,v)$, $\mathbf{p}^1(\omega(z,v))$ and $\mathbf{p}^1(\omega^2(z,v))$. Moreover, we endow the geodesic $p^1(z,v)$ with the induced orientation, so that the endpoints $p(z,v)$ and $\mathbf{p}(z,\omega v)$ correspond to $-\infty$ and $+\infty$, respectively.

We now define the other three operations on $T^1\mathbf{H}$. Let $(z,v) \in T^1\mathbf{H}$. Let $f_{(z,v)} \in \mathbf{Mob}(\mathbf{H})$ be the rotation of order two about the point $\text{foot}_{\mathbf{p}^1(z,v)}(z) \in \mathbf{p}^1(z,v)$. Set $\mathcal{R}(z,v) = f(z,v)$. Clearly the map $\mathcal{R} : T^1\mathbf{H} \to T^1\mathbf{H}$ is a diffeomorphism of order two. Also the map $\mathcal{R}$ commutes with every element from $\mathbf{Mob}(\mathbf{H})$ and the induced diffeomorphism $\mathcal{R} : T^1S \to T^1S$ is well defined. Note that $\mathbf{p}(z,v) = \mathbf{p}(\omega(\mathcal{R}(z,v)))$ and $\mathbf{p}^1(z,v) = \mathbf{p}^1(\mathcal{R}(z,v))$.

Remark. The group $\langle \omega, \mathcal{R}, \rangle$ generated by $\omega, \mathcal{R} : T^1S \to T^1S$ is isomorphic to $\mathbf{Z}_2 \star \mathbf{Z}_3$. It is easy to see that there exists a finite orbit under the action of this group if and only if the surface S is modular, that is S is isomorphic to $\mathbf{H}/G$ where G is a finite index subgroup of $\mathbf{PSL}(2,\mathbf{Z})$. This indicates the relevance of this group action to the Ehrenpreis conjecture.

Let $(z,v) \in T^1\mathbf{H}$, $w \in \mathbf{p}^1(z,v)$, and $t \in \mathbf{R}$. Let $f_{(z,v)}(t) \in \mathbf{Mob}(\mathbf{H})$ be the hyperbolic transformation which fixes the points $\mathbf{p}(z,v)$ and $\mathbf{p}(z,\omega v)$, and so that the signed hyperbolic distance between the points w and $f_{(z,v)}(t)(w)$, is equal to t. Here the signed distance between w and $f_{(z,v)}(t)(w)$ is positive if and only if the points $\mathbf{p}(z,v)$, w, $f_{(z,v)}(t)(w)$, and $\mathbf{p}(z,\omega v)$, sit on the geodesic $\mathbf{p}^1(z,v)$, in this order, with respect to the induced orientation on $\mathbf{p}^1(z,v)$. The definition of the map $f_{(z,v)}(t)$ does not depend on the choice of $w \in \mathbf{p}^1(z,v)$. Clearly, the collection of transformations $f_{(z,v)}(t)$, $t \in \mathbf{R}$, is an one-parameter Abelian group. Set $\mathcal{S}_t(z,v) = f_{(z,v)}(t)(z,v)$. The collections of maps $\mathcal{S}_t : T^1\mathbf{H} \to T^1\mathbf{H}$, $t \in \mathbf{R}$, is an one-parameter Abelin group of diffeomorphisms, and $\mathcal{S}_t$ commutes with every element from $\mathbf{Mob}(\mathbf{H})$. Note that $\mathbf{p}(z,v) = \mathbf{p}(\mathcal{S}_t(z,v))$, and $\mathbf{p}^1(z,v) = \mathbf{p}^1(\mathcal{S}_t(z,v))$.

Remark. In fact, $\mathcal{S}_t : T^1\mathbf{H} \to T^1\mathbf{H}$, is the "equidistant" flow. This means that for $(z_t, v_t) = \mathcal{S}_t(z,v)$, the points z_t move along the line that is equidistant from the geodesic $\mathbf{p}^1(z,v)$ (the distance between z_t and $\mathbf{p}^1(z,v)$ is $\log\sqrt{3}$).

Let $(z,v) \in T^1\mathbf{H}$ so that $\mathbf{p}(z,v) \neq \infty$. Let $\gamma \in \Gamma(\mathbf{H})$, be the geodesic that connects $\mathbf{p}(z,v)$ and ∞, and let $f_\gamma : \mathbf{H} \to \mathbf{H}$, be the reflection through γ. Set $\mathcal{I}^L_1(z,v) = f_\gamma(z,v)$, and $\mathcal{I}^L = \omega^2 \circ \mathcal{I}^L_1$. The maps $\mathcal{I}^L_1, \mathcal{I}^L : T^1\mathbf{H} \to T^1\mathbf{H}$ are defined almost everywhere on $T^1\mathbf{H}$, because the subset of $T^1\mathbf{H}$ on which $\mathbf{p}(z,v) = \infty$ has zero measure in $T^1\mathbf{H}$. The maps $\mathcal{I}^L$ and $\mathcal{I}^L_1$ commute with every element of $\mathbf{Mob}_\infty(\mathbf{H})$. Note that $\mathcal{I}^L_1$ is of order two. Also, $\mathbf{p}(z,v) = \mathbf{p}(\omega(\mathcal{I}^L(z,v)))$. We set $\mathcal{I}^R = (\mathcal{I}^L)^{-1}$.

Remark. The map $\mathcal{I}^L : T^1\mathbf{H} \to T^1\mathbf{H}$ does not commute with the entire group $\mathbf{Mob}(\mathbf{H})$ and the map $\mathcal{I}^L$ does not give rise to a map on T^1S. Let $c \in \operatorname{Cusp}(G)$, and let $G = G_c$ be normalised. Then $\mathcal{I}^L$ commutes with the translation for 1, so the maps $\mathcal{I}^L$ and $\mathcal{I}^L_1$, are well defined almost everywhere on $T^1\mathcal{H}_c(0)$. We have the induced maps $\mathcal{I}^L, \mathcal{I}^L_1 : T^1\mathbf{H} \setminus T^1\mathbf{Th}_G(0) \to T^1\mathbf{H} \setminus T^1\mathbf{Th}_G(0)$, and $\mathcal{I}^L, \mathcal{I}^L_1 : T^1S \setminus T^1\mathbf{Th}_S(0) \to T^1S \setminus T^1\mathbf{Th}_S(0)$, that are defined on each 0-horoball in the normalised setting. Also, note that $\mathcal{I}^L \neq \mathcal{I}^R$.

Proposition 5.1. *We have* $\operatorname{Jac}(\mathcal{R}) = \operatorname{Jac}(\mathcal{S}_t) = 1$ *everywhere on* $T^1\mathbf{H}$ *and* $\operatorname{Jac}(\mathcal{I}^L) = 1$ *almost everywhere on* $T^1\mathbf{H}$*. Moreover we have that the relations* $\mathcal{R} \circ \mathcal{S}_t = \mathcal{S}_{(-t)} \circ \mathcal{R}$ *and* $\mathcal{I}^L \circ \mathcal{S}_t = \mathcal{S}_{(-t)} \circ \mathcal{I}^L$ *hold for every* $t \in \mathbf{R}$*. Also* $\mathcal{R}^2$ *is the identity mapping on* $T^1\mathbf{H}$*.*

Proof. We have already observed that $\mathcal{R}$ is of order two, that is $\mathcal{R}^2$ is the identity mapping on $T^1\mathbf{H}$. If $(z,v),(z_1,v_1) \in T^1\mathbf{H}$, and $(z_1,v_1) = \mathcal{S}_t(z,v)$, for some $t \in \mathbf{R}$, then $\mathbf{p}^1(z,v) = \mathbf{p}^1(z_1,v_1)$, and $\mathbf{p}^1(\mathcal{I}^L(z,v)) = \mathbf{p}^1(\mathcal{I}^L(z_1,v_1))$. This yields the relation $\mathcal{I}^L \circ \mathcal{S}_t = \mathcal{S}_{(-t)} \circ \mathcal{I}^L$. The relation $\mathcal{R} \circ \mathcal{S}_t = \mathcal{S}_{(-t)} \circ \mathcal{R}$, is proved similarly.

Since all four maps $\mathcal{R}$, $\mathcal{S}_t$, $\mathcal{I}^L_1$, and $\mathcal{I}^L$ commute with $\mathbf{Mob}_\infty(\mathbf{H})$, and $\mathbf{Mob}_\infty(\mathbf{H})$ acts transitively on $\mathbf{H}$, we have that the functions $\operatorname{Jac}(\mathcal{R})$, $\operatorname{Jac}(\mathcal{S}_t)$, $\operatorname{Jac}(\mathcal{I}^L_1)$, and $\operatorname{Jac}(\mathcal{I}^L)$, are constant functions almost everywhere on $T^1\mathbf{H}$. Since $\mathcal{R} = \mathcal{R}^{-1}$ we have that $\operatorname{Jac}(\mathcal{R}) = 1$ everywhere on $T^1\mathbf{H}$. It follows from $\mathcal{R} \circ \mathcal{S}_t = \mathcal{S}_{(-t)} \circ \mathcal{R}$, that $\operatorname{Jac}(\mathcal{S}_t) = \operatorname{Jac}(\mathcal{S}_{(-t)})$. Since $(\mathcal{S}_t)^{-1} = \mathcal{S}_{(-t)}$, we conclude that $\operatorname{Jac}(\mathcal{S}_t) = 1$ everywhere on $T^1\mathbf{H}$. From $(\mathcal{I}^L_1)^{-1} = \mathcal{I}^L_1$, we find that $\operatorname{Jac}(\mathcal{I}^L_1) = 1$. Since $\operatorname{Jac}(\omega) = 1$ it follows from the definition of $\mathcal{I}^L$ that $\operatorname{Jac}(\mathcal{I}^L) = 1$. □

Definition 5.2. *Let* $(z,v) \in T^1\mathbf{H}$*, and let* $p \in \partial\mathbf{H}$*.*

- *Denote by* $\Theta_p(z,v)$ *the unique number in* $[-\pi,\pi)$ *that is equal to the positively oriented angle between the vector* v *and the geodesic ray* γ^p_z*, where* γ^p_z *denotes the geodesic ray that starts at* z *and ends at* p*.*
- *By* $\mathcal{H}^p_z$ *we denote the unique horoball that contains* z *and meets* $\partial\mathbf{H}$ *at* p*.*
- *Let* $z,z' \in \mathbf{H}$*. Denote by* $\Delta_p[z,z']$ *the signed hyperbolic distance between the horocircles* $\partial\mathcal{H}^p_z$ *and* $\partial\mathcal{H}^p_{z'}$*. That is,* $\Delta_p[z,z']$ *is non-negative if and only*

if $\mathcal{H}^p_z$ is contained in $\mathcal{H}^p_{z'}$ (in this case we say that z is closer to p than z' is).

- *Suppose that $p \neq \mathbf{p}(z,v)$. Then by $\mathbf{z}_{\max}(\gamma_{(z,v)},p)$ we denote the point on $\gamma_{(z,v)}$ that is the closest to p.*

Also, if $p \neq \mathbf{p}(z,v)$, then there exists a unique point on $\gamma_{(z,v)}$ that is the closest to p. This shows that $\mathbf{z}_{\max}(\gamma_{(z,v)},p)$ is well defined. If $p = \mathbf{p}(z,v)$, then such a point does not exist.

Note that for $z, w \in \mathbf{H}$, we have $\Delta_p[z,w] = -\Delta_p[w,z]$, and $|\Delta_p[z,w]| \leq \mathbf{d}(z,w)$. For points $z, w_1, w_2 \in \mathbf{H}$, we have $\Delta_p[z,w_2] - \Delta_p[z,w_1] = \Delta_p[w_1,w_2]$.

Remark. Let $c \in \mathrm{Cusp}(G)$, and $z, z' \in \mathbf{H}$. Then $\Delta_c[z,z']$ measures the difference in heights, that is

$$\mathbf{h}_c(z) - \mathbf{h}_c(z') = \Delta_c[z,z'].$$

It follows from the definition that $\mathbf{z}_{\max}(\gamma_{(z,v)},p) \neq z$, if and only if $0 < |\Theta_p(z,v)| < \frac{\pi}{2}$. If $\Theta_p(z,v) = 0$, then $\gamma^p_z = \gamma_{(z,v)}$. If $0 < |\Theta_p(z,v)| < \frac{\pi}{2}$, then there exists $t_0 > 0$, so that $\gamma_{(z,v)}(t_0) = \mathbf{z}_{\max}(\gamma_{(z,v)},p)$. Most calculations in the remainder of this paper are based on the following two elementary identities

$$\Delta_p[\mathbf{z}_{\max}(\gamma_{(z,v)},p),z] = \log(\csc(|\Theta_p(z,v)|)), \tag{28}$$

and

$$t_0 = \log\big(\csc(|\Theta_p(z,v)|) + \cot(|\Theta_p(z,v)|)\big). \tag{29}$$

It follows that

$$0 < t_0 - \Delta_p[\mathbf{z}_{\max}(\gamma_{(z,v)},p),z] < \log 2, \tag{30}$$

and when $|\Theta_p(z,v)|$ is small, we have

$$0 < t_0 - \Delta_p[\mathbf{z}_{\max}(\gamma_{(z,v)},p),z] = \log 2 - O(|\Theta_p(z,v)|^2). \tag{31}$$

Proposition 5.2. *Let $(z,v) \in T^1\mathbf{H}$, and $p \in \partial\mathbf{H}$, such that $0 < |\Theta_p(z,v)| < \frac{\pi}{2}$. Let $t_0 > 0$ be the number so that $\mathbf{z}_{\max}(\gamma_{(z,v)},p) = \gamma_{(z,v)}(t_0)$. Then for every $0 \leq t \leq t_0$, we have*

$$t - \log 2 < \Delta_p[\gamma_{(z,v)}(t)),z] < t.$$

Also,

$$e^{-t_0} < |\Theta_p(z,v)| < \pi e^{-t_0}.$$

Proof. The first inequality follows from (30). The second inequality follows from (28), and the fact that for $|\theta| \leq \frac{\pi}{2}$, we have $|\sin\theta| \leq |\theta| \leq \frac{\pi}{2}|\sin\theta|$. □

Proposition 5.3. *Let $t \in \mathbf{R}$, $p \in \partial\mathbf{H}$, and $(z,v) \in T^1\mathbf{H}$, such that $\Theta_p(z,\omega^2 v) = 0$. Let $(z_t,v_t) = \mathcal{S}_t(z,v)$. Then*

$$|\Delta_p[z,z_t]| \geq |t| - \log 6.$$

Proof. Let $w = \mathrm{foot}_{\mathbf{p}^1(z,v)}(z)$, and $w_t = \mathrm{foot}_{\mathbf{p}^1(z,v)}(z_t)$. Since $\Theta_p(z,\omega^2 v) = 0$, we have $w = \mathbf{z}_{\max}(\mathbf{p}^1(z,v),p)$. It follows from the definition of $\mathcal{S}_t(z,v)$, that the hyperbolic distance between the points w and w_t, is equal to $|t|$. From the first inequality in the previous proposition we have

$$|\Delta_p[w,w_t]| \geq |t| - \log 2.$$

Since $\mathbf{d}(z,w)=\mathbf{d}(z_t,w_t)=\log\sqrt{3}$, we have

$$|\Delta_p[z,z_t]|\geq|\Delta_p[w,w_t]|-\log 3\geq|t|-\log 6.$$

If p is an endpoint of $\mathbf{p}^1(z,v)$, then $|\Delta_p[z,z_t]|=|t|>|t|-\log 6$. $\square$

Proposition 5.4. *Let $(z,v)\in T^1\mathbf{H}$, and let $p_i\in\partial\mathbf{H}$, $i=0,1,2$, so that $|\Theta_{p_i}(z,\omega^i v)|<\frac{\pi}{2}\delta$, for some $0\leq\delta$. Let $T\in\mathcal{T}(\mathbf{H})$ be the triangle with the vertices p_i. There exists a universal constant $C>0$, so that for δ small enough, we have $\mathbf{d}(\operatorname{ct}(T),z)\leq C\delta$.*

Proof. We consider the unit disc model $\mathbf{D}$ for $\mathbf{H}$. We may assume that $z=0$, and that $v=\frac{\partial}{\partial x}(0)$. Then $\mathbf{p}(z,\omega^i v)=\omega^i=e^{2\pi i/3}\in\partial\mathbf{D}$, $i=0,1,2$, where $\partial\mathbf{D}$ is the unit circle. Since $|\Theta_{p_i}(z,\omega^i v)|<\frac{\pi}{2}\delta$, we have that $|p_i-\mathbf{p}(z,\omega^i v)|<\frac{\delta}{4}$, where $|p_i-\mathbf{p}(z,\omega^i v)|$ is the Euclidean distance. Recall that z is the centre of the triangle $\mathbf{p}^2(z,v)$. The centre of an ideal triangle, as the function of triples of points on $\partial\mathbf{D}$, is smooth in some neighbourhood the triple $(1,e^{2\pi/3},e^{4\pi/3})$. Therefore, there exists a universal constant $C>0$, so that $\mathbf{d}(\operatorname{ct}(T),z)\leq C\delta$, for δ small. $\square$

Proposition 5.5. *Let $(z,v)\in T^1\mathbf{H}$, and set $\mathcal{R}(z,v)=(z_1,v_1)$. Let $p_i\in\partial\mathbf{H}$, $i=0,1$, so that $|\Theta_{p_i}(z,\omega^i v)|<\frac{\pi}{2}\delta$, for some $0\leq\delta$. By $\gamma\in\Gamma(\mathbf{H})$, we denote the geodesic with the endpoints p_0 and p_1. There exists a universal constant $C>0$, so that for δ small enough we have $\mathbf{d}(\operatorname{foot}_{\mathbf{p}^1(z,v)}(z),\operatorname{foot}_\gamma(z))\leq C\delta$, $\mathbf{d}(\operatorname{foot}_{\mathbf{p}^1(z,v)}(z_1),\operatorname{foot}_\gamma(z_1))\leq C\delta$, and $\mathbf{d}(\operatorname{foot}_\gamma(z),\operatorname{foot}_\gamma(z_1))\leq C\delta$.*

Proof. We consider the unit disc model $\mathbf{D}$ for $\mathbf{H}$. We may assume that the geodesic $\mathbf{p}^1(z,v)$ connects the points -1 and 1, on $\partial\mathbf{D}$, and that $\mathbf{p}(z,v)=-1$. Also, we may assume that the x-coordinate of z is equal to zero. Then $z=i\frac{\sqrt{3}-1}{\sqrt{3}+1}$, and $z_1=-i\frac{\sqrt{3}-1}{\sqrt{3}+1}$. Since $|\Theta_{p_i}(z,\omega^i v)|<\frac{\pi}{2}\delta$, $i=0,1$, we have that for some universal constant $D>0$, and for δ small enough, the inequalities $|p_0+1|<D\delta$, and $|p_1-1|<D\delta$, hold. Here $|p_0+1|$, and $|p_1-1|$, are the Euclidean distances. The function $\operatorname{foot}_\gamma(z)$, as a function of p_0, and p_1, is smooth in some neighbourhood of the pair $(1,-1)$. This shows that there is a universal constant $C>0$, so that for δ small enough, we have $\mathbf{d}(\operatorname{foot}_{\mathbf{p}^1(z,v)}(z),\operatorname{foot}_\gamma(z)),\mathbf{d}(\operatorname{foot}_{\mathbf{p}^1(z,v)}(z_1),\operatorname{foot}_\gamma(z_1)),\mathbf{d}(\operatorname{foot}_\gamma(z),\operatorname{foot}_\gamma(z_1))\leq C\delta$. $\square$

5.2. **The Absorption maps.** First we give a short overview of the construction that follows. In order to construct the measures from the statement of Theorem 4.1, for every $r>2$ we define a map (almost everywhere on $T^1\mathbf{H}$ with respect to the Liouville measure Λ on $T^1\mathbf{H}$)

$$\mathbf{a}_r:T^1\mathbf{H}\to\operatorname{Cusp}(G),$$

such that for almost every $(z,v)\in T^1\mathbf{H}$ we have

- $\mathbf{a}_r(g(z,v))=g(\mathbf{a}_r(z,v))$, for every $g\in G$.
- $\Theta_{\mathbf{a}_r(z,v)}(z,v)\leq Ce^{-r}$.

Moreover we have the induced map $\mathbf{a}^1_r:T^1\mathbf{H}\to\Gamma(G)$ where $\mathbf{a}^1_r(z,v)$ is the geodesic with the endpoints $\mathbf{a}_r(z,v)$ and $\mathbf{a}_r(z,\omega v)$. Also we have the map $\mathbf{a}^2_r:T^1\mathbf{H}\to\mathcal{T}(G)$ where $\mathbf{a}^2_r(z,v)$ is the triangle with vertices $\mathbf{a}_r(z,\omega^j v$, $j=0,1,2$. The maps $\mathbf{a}_r$, $\mathbf{a}^1_r$ and $\mathbf{a}^2_r$ are G equivariant so we have the induced maps from T^1S to $\operatorname{Cusp}(G)$, $\Gamma(S)$ and $\mathcal{T}(S)$, respectively. Note that if $\varphi:S\to\mathbf{R}$ is a non-negative integrable function on S, then $(\mathbf{a}_r^2)_*(\varphi\, d\Lambda) \in \mathcal{M}(\mathcal{T}(S))$. The measure $\mu \in \mathcal{M}(\mathcal{T}(S))$ from Theorem 4.1 will be constructed in this way for a suitable choice of φ.

Our goal is to let r be large and show that for "most" points $(z,v) \in T^1\mathbf{H}$ we have $\mathbf{a}_r^1(z,v) = \mathbf{a}_r^1(\mathcal{R}(z,v))$. This would imply that $\widehat{\partial}(\mathbf{a}_r^2)_*(d\Lambda) = \alpha + \beta$, where α is Ce^{-r} symmetric and β is small. This seems to be a good candidate for the choice of measure μ in Theorem 4.1. However with this choice we have

$$\int\limits_{N^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1) d\widehat{\partial}(\mathbf{a}_r^2)_*(d\Lambda) = \infty,$$

because of the thin part of S. In order to overcome this problem we set $\varphi_r(z) = e^{\frac{1}{2}(2r - \mathbf{h}(z))}$, and consider the measure $(\mathbf{a}_r^2)_*(\varphi_r\, d\Lambda) \in \mathcal{M}(\mathcal{T}(S))$ (note that $\varphi_r(z) = 1$ for $z \in \mathbf{Th}_S(2r)$). But this introduces an imbalance, that is the measure $\widehat{\partial}(\mathbf{a}_r^2)_*(\varphi_r\, d\Lambda)$ is no longer almost symmetric outside $\mathbf{Th}_S(2r)$ because $\varphi_r(z,v) \neq \varphi_r(\mathcal{R}(z,v))$ for z that is is outside $\mathbf{Th}_S(2r)$. That is why we introduce the map

$$\mathbf{b}_r^2 : \mathcal{C}_r(S) \to \mathcal{T}(S),$$

where $\mathcal{C}_r(S) \subset (T^1S \setminus \mathbf{Th}_S(2r))$ is the "correctable" part. Then we let our measure $\mu \in \mathcal{M}(\mathcal{T}(S))$ be given by

$$\mu = (\mathbf{a}_r^2)_*(\varphi_r\, d\Lambda) + 3(\mathbf{b}_r^2)_*(\vartheta_r\, d\Lambda),$$

where $\vartheta_r(z,v) = \varphi_r(\mathcal{R}(z,v)) - \varphi_r(z,v)$, for $(z,v) \in \mathcal{C}_r(S)$. Then we show that the measure $\widehat{\partial}\mu$ has the desired decomposition.

We now return to the construction. Recall that S is a fixed surface of type $(\mathbf{g}, \mathbf{n})$, and $\text{Cusp}(S) = \{c_1(S), ..., c_n(S)\}$. As always, G is one of the n normalised Fuchsian groups $G_{c_i(S)}$, $i = 1, ..., n$, such that $\mathbf{H}/G$ is isomorphic to S. Recall that the interiors of the different 0-horoballs on S are disjoint. This implies that different 1-horoball are disjoint on S

Let $(z,v) \in T^1\mathbf{H}$. For any $t \geq 0$, the point $\gamma_{(z,v)}(t)$ is either in the interior of $\mathbf{Th}_G(1)$, or in one of the 1-horoballs. Let $(c_1, c_2, ...)$, be the ordered set of cusps from $\text{Cusp}(G)$, so that the ray $\gamma_{(z,v)}$ intersects each horoball $\mathcal{H}_{c_i}(1)$, and which are ordered so that the ray $\gamma_{(z,v)}$ visits the horoball $\mathcal{H}_{c_i}(1)$ before it visits the horoball $\mathcal{H}_{c_j}(1)$ if and only if $i < j$. In other words, let $t_i \geq 0$, so that $\gamma_{(z,v)}(t_i) = \mathbf{z}_{\max}(\gamma_{(z,v)}, c_i)$. Then $i < j$, if and only if $t_i < t_j$. Since the 1-horoballs are disjoint, and each horoball is a convex subset of $\mathbf{H}$, we have that $c_i \neq c_j$, for $i \neq j$.

Remark. For a point $(z,v) \in T^1\mathbf{H}$, the set $(c_1, c_2 ...)$ is either: empty, finite but non-empty, or infinite. It can be show that the set of points in $T^1\mathbf{H}$, for which the corresponding set $(c_1, c_2 ...)$ is infinite, has the full measure in $T^1\mathbf{H}$ (we do not use this result).

We now define the maps $\mathbf{a}_r$, and $\mathbf{a}_r^i$, $i = 1, 2$, described above. Given $r > 0$, to every point $(z,v) \in T^1S$, we associate a triangle from $\mathcal{T}(S)$.

Definition 5.3. *Fix $r > 2$. Let $(z,v) \in T^1\mathbf{H}$, and let $(c_1, c_2, ...)$, be the corresponding ordered set of cusps that $\gamma_{(z,v)}$ intersects.*

- *Define $\mathbf{a}_r(z,v) = c_i$, if*

$$\mathbf{h}_{c_i}(\mathbf{z}_{\max}(\gamma_{(z,v)}, c_i)) - \mathbf{h}_{c_i}(z) \geq r, \tag{32}$$

and if this inequality does not hold for any cusp c_j, where $j < i$.

- *Let $0 \leq \mathbf{t}_r(z,v) \leq \infty$, be such that $\gamma_{(z,v)}(\mathbf{t}_r(z,v))$ is the first point of entry of the ray $\gamma_{(z,v)}$ in the horoball $\mathcal{H}_{\mathbf{a}_r(z,v)}(1)$. If (z,v) does not get absorbed, then $\mathbf{t}_r(z,v) = \infty$.*
- *If both $\mathbf{a}_r(z,v)$ and $\mathbf{a}_r(z,\omega v)$ are well defined, then by $\mathbf{a}_r^1(z,v) \in \Gamma(G)$, we denote the geodesic with the endpoints $\mathbf{a}_r(z,v)$ and $\mathbf{a}_r(z,\omega v)$. If all three cusps $\mathbf{a}_r(z,v)$, $\mathbf{a}_r(z,\omega v)$, and $\mathbf{a}_r(z,\omega^2 v)$, are well defined, then by $\mathbf{a}_r^2(z,v) \in \mathcal{T}(G)$, we denote the triangle with the vertices $\mathbf{a}_r(z,v)$, $\mathbf{a}_r(z,\omega v)$, and $\mathbf{a}_r(z,\omega^2 v)$.*

Since $r > 2$, we have that $|\Theta_{\mathbf{a}_r(z,v)}(z,v)| \leq \pi e^{-2} < \frac{\pi}{6}$. This implies that $|\Theta_{\mathbf{a}_r(z,\omega v)}(z,v)| > \frac{\pi}{2}$, so $\mathbf{a}_r(z,v) \neq \mathbf{a}_r(z,\omega v) \neq \mathbf{a}_r(z,\omega^2 v) \neq \mathbf{a}_r(z,v)$. This shows that $\mathbf{a}_r^1(z,v)$ and $\mathbf{a}_r^2(z,v)$ are well defined.

Remark. It is useful to pause here and observe that if $c \in \operatorname{Cusp}(G)$, is such that $\mathbf{a}_r(z,v) = c$, then $|\Theta_c(z,v)| \leq \pi e^{-r}$. This follows from Proposition 5.2. We will often use this observation in the arguments that follow.

We will see below that the function $\mathbf{a}_r : T^1\mathbf{H} \to \operatorname{Cusp}(G)$ is defined almost everywhere on $T^1\mathbf{H}$. Let $g \in G$. We have $g(\mathbf{a}_r(z,v)) = \mathbf{a}_r(g(z,v))$. This shows that the set where $\mathbf{a}_r$ is defined is invariant under G. The induced map $\mathbf{a}_r : T^1S \to \operatorname{Cusp}(S)$, is also denoted by $\mathbf{a}_r$. The absorption time $\mathbf{t}_r(z,v)$ is defined in the same way. The maps $\mathbf{a}_r^1 : T^1S \to \Gamma(S)$, and $\mathbf{a}_r^2 : T^1S \to \mathcal{T}(S)$, are defined accordingly.

Proposition 5.6. *Fix $r > 2$, and $(z,v) \in T^1\mathbf{H}$. Then for every $3r + \max\{0, \mathbf{h}(z)\} \leq t < \mathbf{t}_r(z,v)$, we have $\gamma_{(z,v)}(t) \in \mathbf{Th}_G(1)$. Moreover, suppose that $z \in \mathcal{H}_{c_1}(r+2)$, for some $c_1 \in \operatorname{Cusp}(G)$, and that $\mathbf{a}_r(z,v) \neq c_1$. If $c_2 \in \operatorname{Cusp}(G)$ is the the second cusp (after c_1) so that $\gamma_{(z,v)}$ visits the horoball $\mathcal{H}_{c_2}(1)$, then $\mathbf{a}_r(z,v) = c_2$.*

Proof. For each cusp $c_i \in \operatorname{Cusp}(G)$, such that $\gamma_{(z,v)}$ enters this cusp, we let $t_{\text{ent}}(i) \geq 0$, so that $\gamma_{(z,v)}(t_{\text{ent}}(i))$ is the point of entry of the ray $\gamma_{(z,v)}$ in $\mathcal{H}_{c_i}(1)$. By $t_{\text{ex}}(i) > 0$, we denote the exit time, that is $\gamma_{(z,v)}(t_{\text{ex}}(i))$ is the exit point of the ray $\gamma_{(z,v)}$ from $\mathcal{H}_{c_i}(1)$. If $z \in \mathcal{H}_{c_1}(1)$, then $t_{\text{ent}}(1) = 0$. If $\gamma_{(z,v)}$ never leaves some $\mathcal{H}_{c_i}(1)$, then $t_{\text{ex}}(i) = \infty$.

It follows from Proposition 5.2 that if $t_{\text{ent}}(i) > 0$, then

$$\Delta_{c_i}[\mathbf{z}_{\max}(\gamma_{(z,v)}, c_i), z] \geq \frac{t_{\text{ex}}(i) + t_{\text{ent}}(i)}{2} - \log 2. \tag{33}$$

If $z \in \mathcal{H}_{c_1}(1)$, then

$$\Delta_{c_i}[\mathbf{z}_{\max}(\gamma_{(z,v)}, c_i), z] \geq \frac{t_{\text{ex}}(i) + \mathbf{h}_{c_i}(z)}{2} - \log 2. \tag{34}$$

If $t < \mathbf{t}_r(z,v)$, and $\gamma_{(z,v)}(t) \in \mathcal{H}_{c_i}(1)$, then either $t_{\text{ent}}(i) > 0$, or $i = 1$ and $z \in \mathcal{H}_{c_1}(1)$. In either case, we have $\mathbf{t}_r(z,v) > t_{\text{ex}}(i)$, since $\mathbf{a}_r(z,v) \neq c_i$ (the identity $\mathbf{a}_r(z,v) = c_i$ would contradict the assumption $t < \mathbf{t}_r(z,v)$). In the first case, from (33) we obtain $t \leq t_{\text{ex}} \leq 2r + 2\log 2 < 3r$. In the second case, from (34) we have $t \leq t_{\text{ex}}(1) \leq 2r + \mathbf{h}_{c_1}(z) + 2\log 2 < 3r + \mathbf{h}_{c_1}(z)$. This proves the first part of the proposition.

Assume that $\mathbf{h}_{c_1}(z) \geq r+2$. Let $c_2 \in \operatorname{Cusp}(G)$, be the cusp so that the horoball $\mathcal{H}_{c_2}(1)$ is the second 1-horoball that the ray $\gamma_{(z,v)}$ enters. By (33) we have

$$\mathbf{h}_{c_2}(\mathbf{z}_{\max}(\gamma_{(z,v)}, c_2)) - \mathbf{h}_{c_2}(z) \geq t_{\text{ent}}(2) - \log 2 \geq \mathbf{h}_{c_1}(z) - 1 - \log .$$

Since $\mathbf{h}_{c_1}(z) \geq r+2$, and since $\mathbf{a}_r(z,v) \neq c_1$, we conclude $\mathbf{a}_r(z,v) = c_2$.

□

We want to show that the absorption map is well defined almost everywhere on $T^1\mathbf{H}$ (and on T^1S). We saw in the previous proposition that the ray $\gamma_{(z,v)}$, will get absorbed if it leaves $\mathbf{Th}_S(1)$ after the time $3r + |\mathbf{h}(z)|$. In order to show that $\mathbf{a}_r$ is defined almost everywhere, we need to estimate the Liouville measure of the geodesics segments of a given length that stay in $\mathbf{Th}_S(1)$. The following lemma is probably known, but we prove it in the appendix.

Lemma 5.1. *Let $t > 0$, and let $A(t) \subset T^1\mathbf{Th}_S(1)$ be such that $(z,v) \in A(t)$ if the segment $\gamma_{(z,v)}[0,t]$ is contained in $\mathbf{Th}_S(1)$. Then there are constants $C(S), q(S) > 0$, that depend only on S, such that $\Lambda(A(t)) < C(S)e^{-q(S)t}$.*

We have

Proposition 5.7. *Fix $r > 2$. The map $\mathbf{a}_r : T^1S \to \operatorname{Cusp}(S)$, is defined almost everywhere.*

Remark. The proof below uses Lemma 5.1. However one does not need Lemma 5.1 to prove that the absorptions map is defined almost everywhere on $T^1\mathbf{H}$. The fact that the geodesic flow $\mathbf{g}_t$ is ergodic implies that the geodesic ray $\gamma_{(z,v)}$ leaves the thick part $\mathbf{Th}_S(1)$, after the time $3r + |\mathbf{h}(z)|$, for almost every $(z,v) \in T^1S$. Then it follows from Proposition 5.6 that almost every (z,v) will be absorbed. However Lemma 5.1 will be used later in a similar manner, so we decide to state it here and present its first application.

Proof. Let $s > 0$, and let $F_{r,s} \subset T^1\mathbf{Th}_G(s)$ be such that $(z,v) \in F_{r,s}$ if $\mathbf{t}_r(z,v) = \infty$, and $|\mathbf{h}(z)| \leq s$. By Proposition 5.6 we have that $\mathbf{g}_{3r+s}(F_{r,s}) \subset A(t)$, for every $t > 0$ (here $A(t)$ is the set defined in the statement of Lemma 5.1). Since $\Lambda(A(t)) \to 0$ when $t \to \infty$, we find that $\Lambda(\mathbf{g}_{3r+s}(F_{r,s})) = \Lambda(F_{r,s}) = 0$, so the map $\mathbf{a}_r$ is defined almost everywhere on $T^1\mathbf{Th}_G(S)$, for every $s > 0$.

□

Definition 5.4. *For every $r > 2$, we define the map $\mathbf{f}_r : T^1\mathbf{H} \to N^1\Gamma(G)$, as $\mathbf{f}_r(z,v) = \operatorname{foot}_{\mathbf{a}^1_r(z,v)}(\operatorname{ct}(\mathbf{a}^2_r(z,v)))$. That is, $\mathbf{f}_r N^1\Gamma(G)$ is the foot of the centre of the triangle $\mathbf{a}^2_r(z,v) \in \mathcal{T}(G)$.*

The map $\mathbf{f}_r$ commutes with the action of the group G, and we have that the induced map $\mathbf{f}_r : T^1S \to N^1\Gamma(S)$, is well defined.

Proposition 5.8. *There exists a universal constant $C > 0$, and $r_0 > 0$, so that for $r > r_0$, we have that $\mathbf{d}(\mathbf{f}_r(z,v), \operatorname{foot}_{\mathbf{a}^1_r(z,v)}(z)) \leq Ce^{-r}$.*

Proof. It follows from Proposition 5.2 that $|\Theta_{\mathbf{a}_r(z,\omega^j v)}(z,\omega^j v)| \leq \pi e^{-r}$. By Proposition 5.4, for r large enough we have $\mathbf{d}(z, \operatorname{ct}(\mathbf{a}^2_r(z,v))) \leq C\pi e^{-r}$, for some universal constant $C > 0$. This proves the proposition.

□

Denote by $\mathcal{M}_{\Lambda}(T^1S)$, the space of measures from $\mathcal{M}(T^1S)$, that are absolutely continuous with respect to the Liouville measure Λ. Let $\nu \in \mathcal{M}_{\Lambda}(T^1S)$. Then $(\mathbf{a}_r^2)_*(\nu) \in \mathcal{M}(\mathcal{T}(S))$, and $(\mathbf{f}_r)_*(\nu) \in \mathcal{M}(N^1\Gamma(S))$, are well defined since $\mathbf{a}_r^1$ and $\mathbf{a}_r^2$ are defined almost everywhere on T^1S.

Let $\varphi : S \to \mathbf{R}$, be an integrable, non-negative function. We have the induced map $\varphi : T^1S \to \mathbf{R}$, given by $\varphi(z,v) = \varphi(z)$. Then $\varphi\, d\Lambda \in \mathcal{M}_{\Lambda}(T^1S)$. Since $\varphi(z,v) = \varphi(z,\omega v) = \varphi(z,\omega^2 v)$, and since $\mathbf{a}_r^2(z,v) = \mathbf{a}_r^2(z,\omega v) = \mathbf{a}_r^2(z,\omega^2 v)$, we have

$$3(\mathbf{f}_r)_*(\varphi\, d\Lambda) = \widehat{\partial}(\mathbf{a}_r^2)_*(\varphi\, d\Lambda), \tag{35}$$

where $\widehat{\partial} : \mathcal{M}(\mathcal{T}(S)) \to \mathcal{M}(N^1\Gamma(S))$, is the operator defined at the beginning of Section 4.

Let $(z,v) \in T^1\mathbf{H}$, and set $(z_1,v_1) = \mathcal{R}(z,v)$. We say that $(z,v) \in (T^1\mathbf{H})^+$ (or that (z,v) is above the geodesic $\mathbf{p}^1(z,v)$) if $\mathbf{h}(z) > \mathbf{h}(z_1)$ and $(z,v) \in (T^1\mathbf{H})^-$ (or that (z,v) is below the geodesic $(\mathbf{p}^1(z,v))$ if $\mathbf{h}(z) < \mathbf{h}(z_1)$. Since for almost every $(z,v) \in T^1\mathbf{H}$ we have that either $\mathbf{h}(z) > \mathbf{h}(z_1)$ or $\mathbf{h}(z) < \mathbf{h}(z_1)$ we see that $(T^1\mathbf{H})^+ \cup (T^1\mathbf{H})^-$ has full measure in $T^1\mathbf{H}$. The sets $(T^1\mathbf{H})^+$ and $(T^1\mathbf{H})^-$ are disjoint and $\mathcal{R}((T^1\mathbf{H})^-) = (T^1\mathbf{H})^+$. The sets $(T^1\mathbf{H})^+$ and $(T^1\mathbf{H})^-$ are invariant under G and the sets $(T^1S)^+$ and $(T^1S)^-$ are defined accordingly (recall that $\mathbf{h}(z)$ is well defined for $z \in S$).

Definition 5.5. *Fix $r > 2$, and let $(z,v) \in T^1\mathbf{H}$. We say that $(z,v) \in \mathcal{A}_r(G)$ if $(z,v) \in T^1\mathbf{Th}_G(2r) \cap \mathcal{R}\big(T^1\mathbf{Th}_G(2r)\big)$, and if $\mathbf{a}_r^1(z,v) = \mathbf{a}_r^1(\mathcal{R}(z,v))$ (here we assume that both $\mathbf{a}_r^1(z,v)$ and $\mathbf{a}_r^1(\mathcal{R}(z,v))$ are well defined).*

The set $\mathcal{A}_r(G)$ is invariant under G and by $\mathcal{A}_r(S)$ we denote the corresponding subset of T^1S.

Definition 5.6. *Fix $r > 2$ and let $c \in \operatorname{Cusp}(G)$. Let*

$$Q_c = \{(z,v) \in T^1\mathcal{H}_c(2r) \cup \mathcal{R}\big(\mathcal{H}_c(2r)\big) : c \quad \textit{is not an endpoint of} \quad \mathbf{a}_r^1(z,v)\}.$$

Set $\widetilde{Q}_c = Q_c \cap \mathcal{R}(Q_c)$. Let

$$\mathcal{B}_r(G) = \bigcup \{(z,v) \in \widetilde{Q}_c : \mathbf{a}_r^1(z,v) = \mathbf{a}_r^1(z',v'),\ \textit{whenever } (z',v') \in \widetilde{Q}_c \textit{ and } \mathbf{p}^1(z,v) = \mathbf{p}^1(z',v')\},$$

where the union is taken over all $c \in \operatorname{Cusp}(G)$.

The set $\mathcal{B}_r(G)$ is invariant under G and by $\mathcal{B}_r(S)$ we denote the corresponding subset of T^1S. Note that

$$\mathcal{B}_r(G) \subset T^1\mathbf{Thin}_G(2r) \cup \mathcal{R}\big(T^1\mathbf{Thin}_G(2r)\big),$$

so $\mathcal{A}_r(G) \cap \mathcal{B}_r(G) = \emptyset$.

Definition 5.7. *Let $r > 2$ and let $c \in \operatorname{Cusp}(G)$. Let*

$$Q'_c = \{(z,v) \in T^1\mathcal{H}_c(2r) \cap (T^1\mathbf{H})^+ : c \quad \textit{is not an endpoint of} \quad \mathbf{a}_r^1(z,v)\}.$$

Set

$$\mathcal{C}_r(G) = \bigcup_{c \in \operatorname{Cusp}(G)} Q'_c.$$

Let $(z,v) \in \mathcal{C}_r(G)$. We say that $(z,v) \in \mathcal{C}_r^L(G)$, if $\mathbf{a}_r(z,v) = \mathbf{a}_r(\omega(\mathcal{I}^L(z,v)))$, and $(z,v) \in \mathcal{C}_r^R(G)$, if $\mathbf{a}_r(z,\omega v) = \mathbf{a}_r(\mathcal{I}^R(z,v))$.

The sets $\mathcal{C}_r(G)$ and $\mathcal{C}_r^j(G)$ are invariant under G and the corresponding quotients are denoted by $\mathcal{C}_r(S)$ and $\mathcal{C}_r^j(S)$. In the definition of $\mathcal{C}_r^j(G)$, we consider the maps $\mathcal{I}^L, \mathcal{I}^R$, as the maps of $T^1\mathbf{H} \setminus T^1\mathbf{Th}_G(0)$. Since $(z,v) \in (T^1\mathbf{H})^+$, we have that $\mathcal{I}^L(z,v)$, and $\mathcal{I}^R(z,v)$, are well defined. Assume that $(z,v) \in \mathcal{B}_r(G) \cap (T^1\mathbf{H})^+$. Then $z \in \mathcal{H}_c(1)$ for some $c \in \operatorname{Cusp}(G)$ and $\mathbf{h}(z) \geq 2r$ (by definition of $\mathcal{B}_r(G)$). Moreover, neither $\gamma_{(z,v)}$ or $\gamma_{(z,\omega v)}$ gets absorbed by the cusp c. This shows that every such (z,v) belongs to $\mathcal{C}_r(G)$, that is

$$\mathcal{B}_r(G) \cap (T^1\mathbf{H})^+ \subset \mathcal{C}_r(G). \tag{36}$$

Proposition 5.9. *We have $\mathcal{I}^L(\mathcal{C}_r(G)) = \mathcal{C}_r(G)$, and $\mathcal{I}^L(\mathcal{C}_r^L(G)) = \mathcal{C}_r^R(G)$.*

Proof. Let $(z,v) \in \mathcal{C}_r(G)$, and let $(z_1,v_1) = \mathcal{I}^L(z,v)$. There exists a cusp $c \in \operatorname{Cusp}(G)$, so that $z \in \mathcal{H}_c(2r)$. Set $G = G_c$, and $c = \infty$. It follows from the definition of $\mathcal{I}^L$, that $\mathbf{h}_\infty(z) = \mathbf{h}_\infty(z_1)$. Also $\Theta_\infty(z,v) = -\Theta_\infty(z_1,\omega v_1)$, and $\Theta_\infty(z,\omega v) = -\Theta_\infty(z_1,v_1)$. This implies that $\mathbf{h}_\infty(\mathbf{z}_{\max}(\gamma_{(z_1,\omega v_1)},\infty)) = \mathbf{h}_\infty(\mathbf{z}_{\max}(\gamma_{(z,v)},\infty))$, and $\mathbf{h}_\infty(\mathbf{z}_{\max}(\gamma_{(z_1,v_1)},\infty)) = \mathbf{h}_\infty(\mathbf{z}_{\max}(\gamma_{(z,\omega v)},\infty))$.

Since $(z,v) \in \mathcal{C}_r(G)$, we have $\mathbf{a}_r(z,v) \neq \infty \neq \mathbf{a}_r(z,\omega v)$. In order to show $(z_1,v_1) \in \mathcal{C}_r(G)$, we need to show that $\mathbf{a}_r(z_1,v_1) \neq \infty \neq \mathbf{a}_r(z_1,\omega v_1)$. Assume that $\mathbf{a}_r(z_1,v_1) = \infty$. Then $\mathbf{h}_\infty(\mathbf{z}_{\max}(\gamma_{(z_1,v_1)},\infty)) - \mathbf{h}_\infty(z_1) \geq r$. But then $\mathbf{h}_\infty(\mathbf{z}_{\max}(\gamma_{(z,\omega v)},\infty)) - \mathbf{h}_\infty(z) \geq r$, so $\mathbf{a}_r(z,\omega v) = \infty$, which is a contradiction. Similarly we show $\mathbf{a}_r(z_1,\omega v_1) \neq \infty$. Putting this together proves $\mathcal{I}^L(\mathcal{C}_r(G)) \subset \mathcal{C}_r(G)$. In the same way we show $\mathcal{C}_r(G) \subset \mathcal{I}^L(\mathcal{C}_r(G))$.

The equality $\mathcal{I}^L(\mathcal{C}_r^L(G)) = \mathcal{C}_r^R(G)$, follows directly from the definition. □

We are yet to see that $\mathcal{A}_r(S)$, $\mathcal{B}_r(S)$, and $\mathcal{C}_r^j(S)$, are non-empty sets (see Lemma 5.3). The following proposition is elementary and the proof is left to the reader.

Proposition 5.10. *Let $r > 2$. If $(z,v) \in T^1\mathbf{Thin}_G(2r) \cup \mathcal{R}\big(T^1\mathbf{Thin}_G(2r)\big)$ then $0 < 2r - \log 3 < \mathbf{h}(z)$.*

Definition 5.8. *Let $r > 2$, and let $(z,v) \in \mathcal{C}_r(G)$. Let $c \in \operatorname{Cusp}(G)$, so that $z \in \mathcal{H}_c(2r)$. Set $\mathbf{b}_r(z,v) = c$. By $\mathbf{b}_r^1(z,v)$, we denote the geodesic with the endpoints $\mathbf{a}_r(z,v)$ and $\mathbf{b}_r(z,v)$ (providing that $\mathbf{a}_r(z,v)$ exists). By $\mathbf{b}_r^2(z,v) \in \mathcal{T}(G)$, we denote the triangle with the vertices $\mathbf{a}_r(z,v)$, $\mathbf{a}_r(z,\omega v)$, and $\mathbf{b}_r(z,v)$ (providing that $\mathbf{a}_r(z,\omega^j v)$, $j = 0,1$, exist).*

From the definition of the set $\mathcal{C}_r(G)$, we have that $\mathbf{a}_r(z,v) \neq c \neq \mathbf{a}_r(z,\omega v)$ (if $\mathbf{a}_r(z,\omega^i v) = c$, $i = 0,1$, then $\mathbf{t}_r(z,\omega^i v) = 0$). This shows that the maps $\mathbf{b}_r(z,v)$, $\mathbf{b}_r^1(z,v)$, and $\mathbf{b}_r^2(z,v)$ are well defined almost everywhere on $\mathcal{C}_r(G)$. Note that these three maps commute with the action of G, so we have the induced maps $\mathbf{b}_r : \mathcal{C}_r(S) \to \operatorname{Cusp}(S)$, $\mathbf{b}_r^1 : \mathcal{C}_r(S) \to \Gamma(S)$, and $\mathbf{b}_r^2 : \mathcal{C}_r(S) \to \mathcal{T}(S)$. The edges of the triangle $\mathbf{b}_r^2(z,v)$, are $\mathbf{a}_r^1(z,v)$, $\mathbf{b}_r^1(z,v)$, and $\mathbf{b}_r^1(z,\omega v)$.

For $(z,v) \in \mathcal{C}_r(G)$, let $\bar{\mathbf{f}}_r(z,v) = \operatorname{foot}_{\mathbf{a}_r^1(z,v)}(\operatorname{ct}(\mathbf{b}_r^2(z,v)))$, that is $\bar{\mathbf{f}}_r(z,v)$ is the foot of the centre of the triangle $\mathbf{b}_r^2(z,v) \in \mathcal{T}(G)$, with respect to the geodesic $\mathbf{a}_r^1(z,v) \in \Gamma(G)$. Set $\bar{\mathbf{f}}_r^L(z,v) = \operatorname{foot}_{\mathbf{b}_r^1(z,v)}(\operatorname{ct}(\mathbf{b}_r^2(z,v)))$, and $\bar{\mathbf{f}}_r^R(z,v) = \operatorname{foot}_{\mathbf{b}_r^1(z,\omega v)}(\operatorname{ct}(\mathbf{b}_r^2(z,v)))$. The notation $\bar{\mathbf{f}}_r^L(z,v)$ indicates that the point $\bar{\mathbf{f}}_r^L(z,v)$ belongs to the vertical edge of $\mathbf{b}_r^2(z,v)$ that is on the "left-hand" side of $\mathbf{b}_r^2(z,v)$, with the normalisation $G = G_{\mathbf{b}_r(z,v)}$. Similarly, the point $\bar{\mathbf{f}}_r^R(z,v)$ belongs to the vertical edge of $\mathbf{b}_r^2(z,v)$ that is to the "right-hand" side of $\mathbf{b}_r^2(z,v)$.

Proposition 5.11. *There exists $r_0 > 0$, so that for $r > r_0$, the following holds. Let $(z,v) \in \mathcal{C}^L_r(G)$, and set $\mathcal{I}^L(z,v) = (z_1,v_1)$. Assume that $\mathbf{b}^2_r(z,v)$ and $\mathbf{b}^2_r(z_1,v_1)$, exist. Then $\mathbf{d}(\bar{\mathbf{f}}^L_r(z,v), \bar{\mathbf{f}}^R_r(z_1,v_1)) \le e^r$. Similarly, if $(z,v) \in \mathcal{C}^R_r(G)$, for $\mathcal{I}^R(z,v) = (z_1,v_1)$, we have $\mathbf{d}(\bar{\mathbf{f}}^R_r(z,v), \bar{\mathbf{f}}^L_r(z_1,v_1)) \le e^r$, providing that $\mathbf{b}^2_r(z,v)$ and $\mathbf{b}^2_r(z_1,v_1)$, exist.*

Proof. Suppose $(z,v) \in \mathcal{C}^L_r(G)$. Let $c = \mathbf{b}_r(z,v)$, and set $G = G_c$. Let $\mathbf{p}(z_1,v_1) = p_1$, $\mathbf{p}(z,\omega v) = p_2$, and $\mathbf{p}(z_1,\omega v_1) = \mathbf{p}(z,v) = p_3$ (the identity $\mathbf{p}(z_1,\omega v_1) = \mathbf{p}(z,v)$ follows from the definition of $\mathcal{I}^L$). Since $\mathbf{h}(z) = \mathbf{h}_\infty(z) = \mathbf{h}_\infty(z_1) \ge 2r > r+2$, we have by Proposition 5.6 that if $\mathbf{a}_r(z_1,v_1) = c_1 \in \operatorname{Cusp}(G)$, then $\mathcal{H}_{c_1}(1)$, is the first 1-horoball that the geodesic ray $\gamma_{(z_1,v_1)}$ enters after leaving $\mathcal{H}_\infty(1)$. With this normalisation, we have that the Euclidean diameter of the horoball $\mathcal{H}_{c_1}(1)$, is at most 1. Since the ray $\gamma_{(z_1,v_1)}$ ends at the point p_1, we conclude that $|c_1 - p_1| \le 1$, where $|c_1 - p_1|$ is the Euclidean distance between the real numbers c_1 and p_1. Let $\mathbf{a}_r(z,\omega v) = c_2 \in \operatorname{Cusp}(G)$. Similarly, we see that $|c_2 - p_2| \le 1$.

Since $(z,v) \in \mathcal{C}^L_r(G)$, we have $\mathbf{a}_r(z,v) = \mathbf{a}_r(z_1,\omega v_1) = c_3 \in \operatorname{Cusp}(G)$. Similarly we see that $|c_3 - p_3| \le 1$. From the fact that $\mathbf{h}_\infty(z) \ge 2r$, we have that $|p_1 - p_3| = |p_2 - p_3| \ge e^{2r}$. This implies that

$$\log \frac{|c_3 - c_1|}{|c_3 - c_2|} \le \frac{2}{e^{2r}-1} + o\left(\frac{1}{e^{2r}}\right) < e^{-r},$$

for r large enough. Since

$$\mathbf{d}(\bar{\mathbf{f}}^L_r(z,v), \bar{\mathbf{f}}^R_r(z_1,v_1)) = \log \frac{|c_3 - c_1|}{|c_3 - c_2|},$$

the proposition follows. □

Let $\nu \in \mathcal{M}_\Lambda(\mathcal{C}_r(S))$. Then $(\mathbf{b}^2_r)_*(\nu) \in \mathcal{M}(\mathcal{T}(S))$, and $(\bar{\mathbf{f}}_r)_*(\nu), (\bar{\mathbf{f}}^L_r)_*(\nu), (\bar{\mathbf{f}}^R_r)_*(\nu) \in \mathcal{M}(N^1\Gamma(S))$, are well defined since $\mathbf{a}^1_r$ and $\mathbf{a}^2_r$ are defined almost everywhere on $\mathcal{C}_r(S)$. By definition, we have

$$\widehat{\partial}(\mathbf{b}^2_r)_*(\nu) = (\bar{\mathbf{f}}_r)_*(\nu) + (\bar{\mathbf{f}}^L_r)_*(\nu) + (\bar{\mathbf{f}}^R_r)_*(\nu) \tag{37}$$

Definition 5.9. *If $\mathbf{a}_r(z,v)$ is defined, then the r-combinatorial length $\mathcal{K}_r(z,v)$ is defined as follows. Let γ be the geodesic ray that connects z and $\mathbf{a}_r(z,v)$. Let $\iota(\gamma,\tau(G))$, be the number of (transverse) intersections between the ray γ and the edges from $\lambda(G)$. If z does not belong to $\mathcal{H}_{\mathbf{a}_r(z,v)}(1)$, then $\mathcal{K}_r(z,v) = \iota(\gamma,\tau(G))$. If z belongs to $\mathcal{H}_{\mathbf{a}_r(z,v)}(1)$, then $\mathcal{K}_r(z,v) = e^{\mathbf{h}_{\mathbf{a}_r(z,v)}(z)} = e^{\mathbf{h}(z)} > 1$.*

Again, the function $\mathcal{K}_r : T^1S \to \mathbf{R}^+ \cup \{0\}$, is defined almost everywhere.

Proposition 5.12. *There exists $r_0 > 0$, so that for $r > r_0$, the following holds. For $(z,v) \in T^1\mathbf{H}$, we have*

$$\mathcal{K}(\mathbf{a}^1_r(z,v), \mathbf{f}_r(z,v)) \le e(\mathcal{K}_r(z,v) + \mathcal{K}_r(z,\omega v)).$$

Let $(z,v) \in \mathcal{C}_r(G)$. We have

$$\mathcal{K}(\mathbf{a}^1_r(z,v), \bar{\mathbf{f}}_r(z,v)) \le \mathcal{K}_r(z,v) + \mathcal{K}_r(z,\omega v),$$

$$\mathcal{K}(\mathbf{b}^1_r(z,v), \bar{\mathbf{f}}^L_r(z,v)) \le 5(\mathcal{K}_r(z,v) + \mathcal{K}_r(z,\omega v)),$$

and

$$\mathcal{K}(\mathbf{b}^1_r(z,\omega v), \bar{\mathbf{f}}^R_r(z,v)) \le 5(\mathcal{K}_r(z,v) + \mathcal{K}_r(z,\omega v)).$$

Proof. Let γ denote the geodesic ray that connects z and $\mathbf{a}_r(z,v)$, and let γ' denote the geodesic ray that connects z and $\mathbf{a}_r(z,\omega v)$. Then every edge from $\lambda(G)$, that intersects (transversely) the geodesic $\mathbf{a}^1_r(z,v)$, has to intersect one of the rays γ or γ'. This shows

$$\iota(\mathbf{a}^1_r(z,v),\tau(G)) \le \iota(\gamma,\tau(G)) + \iota(\gamma',\tau(G)).$$

Assume that $\mathbf{f}_r(z,v)$ belongs to the 1-horoball of one of the cusps $\mathbf{a}_r(z,v)$ or $\mathbf{a}_r(z,\omega v)$, say $\mathbf{f}_r(z,v)$ belongs to the 1-horoball at the cusp $\mathbf{a}_r(z,v)$. It follows from Proposition 5.8 that for r large enough, we have

$$\mathbf{d}(\mathbf{f}_r(z,v),z) \le \mathbf{d}(\text{foot}_{\mathbf{a}^1_r(z,v)}(z),\mathbf{f}_r(z,v)) + \mathbf{d}(z,\text{foot}_{\mathbf{a}^1_r(z,v)}(z)) \le Ce^{-r} + \log\sqrt{3} < 1.$$

Therefore, we have that $e^{\mathbf{h}_{\mathbf{a}_r(z,v)}(\mathbf{f}_r(z,v))} \le e^{(1+\mathbf{h}_{\mathbf{a}_r(z,v)}(z))}$. This proves the first inequality in this proposition.

Assume now that $(z,v) \in \mathcal{C}_r(G)$. Since $\mathbf{a}_r(z,v) \ne \mathbf{b}_r(z,v) \ne \mathbf{a}_r(z,\omega v)$, we have $\mathcal{K}(\mathbf{a}^1_r(z,v),\bar{\mathbf{f}}_r(z,v)) = \iota(\mathbf{a}^1_r(z,v),\tau(G))$. Let γ and γ' be as above. Then every edge from $\lambda(G)$, that intersects (transversely) the geodesic $\mathbf{a}^1_r(z,v)$, has to intersect one of the rays γ or γ'. This proves the second inequality $\mathcal{K}(\mathbf{a}^1_r(z,v),\bar{\mathbf{f}}_r(z,v)) \le \mathcal{K}_r(z,v) + \mathcal{K}_r(z,\omega v)$.

If we prove

$$\mathcal{K}(\mathbf{b}^1_r(z,v),\bar{\mathbf{f}}^L_r(z,v)) \le 5\mathcal{K}(\mathbf{a}^1_r(z,v),\bar{\mathbf{f}}_r(z,v)),$$

then the third inequality would follow from the second. Note that $\mathbf{d}(\bar{\mathbf{f}}_r(z,v),\bar{\mathbf{f}}^L_r(z,v)) < \log 3$. We have

$$\mathcal{K}(\mathbf{b}^1_r(z,v),\bar{\mathbf{f}}^L_r(z,v)) \le \iota(\mathbf{b}^1_r(z,v),\tau(G)) + e^{h_{\max}(\bar{\mathbf{f}}^L_r(z,v))} \le \iota(\mathbf{b}^1_r(z,v),\tau(G)) + e^{\log 3 + h_{\max}(\bar{\mathbf{f}}_r(z,v))}.$$

The geodesic $\mathbf{a}^1_r(z,v)$ intersects every edge from $\lambda(G)$, that is intersected by $\mathbf{b}^1_r(z,v)$. In addition, the geodesic $\mathbf{a}^1_r(z,v)$ intersects at least $2e^{(h_{\max}(\bar{\mathbf{f}}_r(z,v))-1)}$ edges from $\lambda(G)$, that all have $\mathbf{b}_r(z,v)$ as their endpoints (since $\mathbf{b}^1_r(z,v)$ has $\mathbf{b}_r(z,v)$ as an endpoint, we see that $\mathbf{b}^1_r(z,v)$ can not intersect (transversely) any edge from $\lambda(G)$, that ends at $\mathbf{b}_r(z,v)$). We have

$$\mathcal{K}(\mathbf{a}^1_r(z,v),\bar{\mathbf{f}}_r(z,v)) \ge \iota(\mathbf{b}^1_r(z,v),\tau(G)) + 2e^{(h_{\max}(\bar{\mathbf{f}}_r(z,v))-1)} \ge \frac{2}{3e}\mathcal{K}(\mathbf{b}^1_r(z,v),\bar{\mathbf{f}}^L_r(z,v)),$$

which proves the third inequality. The fourth inequality is proved in the same way. □

5.3. **Certain special sets and their properties.** We have

Definition 5.10. *Let $r > 2$. Define $\varphi_r : S \to \mathbf{R}$, by $\varphi_r(z) = 1$, if $z \in \mathbf{Th}_S(2r)$, and by*

$$\varphi_r(z) = e^{\frac{1}{2}\min\{0,(2r-\mathbf{h}(z))\}}.$$

Note that φ_r is continuous on S, and $\varphi_r \, d\Lambda \in \mathcal{M}_\Lambda(T^1S)$. The induced function $\varphi_r : T^1S \to \mathbf{R}$, is also denoted by φ_r.

Remark. The reason that the factor $\frac{1}{2}$ appears in the definition of φ_r, is Lemma 5.5 below. In fact, we could replace $\frac{1}{2}$, with any number between 0 and 1. Also, note that if $(z,v) \in \mathcal{A}_r(S)$, then $\varphi_r(\mathcal{R}(z,v)) = \varphi_r(z,v) = 1$.

For $r > 2$, the set $\mathcal{D}_r(S) \subset T^1S$, is defined so that $(z, v) \in \mathcal{D}_r(S)$, if $\mathbf{h}(z) \leq 10r$, and if the inequalities $\mathbf{t}_r(z, v), \mathbf{t}_r(z, \omega v), \mathbf{t}_r(z, \omega^2 v) \leq r^2$, hold. Note that if $(z, v) \in \mathcal{D}_r(S)$, then for $r > 4$, we have $\mathbf{h}(T) \leq r^2$, where $T = \mathbf{a}_r^2(z, v) \in \mathcal{T}(S)$ (recall Definition 3.4 for the definition of $\mathbf{h}(T)$). The following lemma will be proved in Section 7.

Lemma 5.2. *There exists $r_0 > 0$, so that for $r > r_0$, we have*

$$\int\limits_{T^1S\setminus\mathcal{D}_r(S)} (\mathcal{K}_r(z, v) + \mathcal{K}_r(z, \omega v) + \mathcal{K}_r(z, \omega^2 v) + 1)\varphi_r(z)\, d\Lambda \leq P(r)e^{-r}.$$

The following lemma will also be proved in Section 7.

Lemma 5.3. *There exists $r_0 > 0$, so that for $r > r_0$, we have*

$$\int\limits_{H_r} (\mathcal{K}_r(z, v) + \mathcal{K}_r(z, \omega v) + \mathcal{K}_r(z, \omega^2 v) + 1)\varphi_r(z)\, d\Lambda \leq P(r)e^{-r},$$

where $H_r = H'_r \cap \mathcal{D}_r(S)$, and

$$H'_r = (T^1S \setminus (\mathcal{A}_r(S) \cup \mathcal{B}_r(S))) \cup (\mathcal{C}_r(S) \setminus \mathcal{C}_r^L(S)) \cup (\mathcal{C}_r(S) \setminus \mathcal{C}_r^R(S)).$$

We now give a better description of the set H_r. The results that follow in this subsection will not be used in the proof of Theorem 4.1 below, so the reader may skip the rest of this subsection and go to the next subsection and the proof of Theorem 4.1.

Proposition 5.13. *Let $z_{-1}, z_1 \in \mathbf{H}$, so that $\mathbf{d}(z_{-1}, z_1) \geq 2$. Let γ be the geodesic that contains z_{-1} and z_1, and assume that $\mathbf{z}_{\max}(\gamma, \infty) = z_0$, where z_0 is the midpoint of the geodesic segment between z_{-1} and z_1. Let $w_{-1}, w_1 \in \mathbf{H}$, so that $\mathbf{d}(z_{-1}, w_{-1}), \mathbf{d}(z_1, w_1) \leq \delta$, for some $0 \leq \delta$, and let γ' be the geodesic that contains w_{-1} and w_1. There exists a universal constant $C > 0$, so that for δ small enough, we have $\mathbf{d}(\mathbf{z}_{\max}(\gamma, \infty), \mathbf{z}_{\max}(\gamma', \infty)) \leq C\delta$.*

Proof. We work in the unit disc $\mathbf{D}$. The point $\infty \in \partial\mathbf{H}$, corresponds to the point $i \in \partial\mathbf{D}$. We may assume that $z_0 = 0$. Then γ is the geodesic that connects -1 and 1. Moreover, $z_{-1} = -x$, and $z_1 = x$, for some $x > (e-1)/(e+1)$. Let q, q_1, be the endpoints of γ'. From the assumption $\mathbf{d}(z_{-1}, w_{-1}), \mathbf{d}(z_1, w_1) \leq \delta$, and since $\mathbf{d}(z_{-1}, z_1) > 2$, we see that there exists a universal constant $D > 0$, so that $|q + 1|, |q_1 - 1| \leq D\delta$. The Möbius transformation f is uniquely determined by the conditions $f(i) = i$, $f(-1) = q$, and $f(1) = q_1$. Moreover, f depends smoothly on q near -1, and q_1 near 1. Therefore, there exists a universal constant $C > 0$, so that $\mathbf{d}(z_0, f(z_0)) = \mathbf{d}(0, f(0)) \leq C\delta$. Since $f(i) = i$, we have $f(z_0) = \mathbf{z}_{\max}(\gamma', i)$, and this proves the proposition. □

Let $z_1, z_2, w_1, w_2 \in \mathbf{H}$, such that $\mathbf{d}(z_1, w_1) = \mathbf{d}(z_2, w_2)$. Let ζ_j, $j = 1, 2$, be the point on the geodesic segment between z_j and w_j, so that $\mathbf{d}(z_1, \zeta_1) = \mathbf{d}(z_2, \zeta_2) = d$. Then $\mathbf{d}(w_1, \zeta_1) = \mathbf{d}(w_2, \zeta_2) = d'$, for some $d' \geq 0$. We have the following elementary inequality in hyperbolic geometry

$$\mathbf{d}(\zeta_1, \zeta_2) \leq D(\mathbf{d}(z_1, z_2)e^{-d} + \mathbf{d}(w_1, w_2)e^{-d'}), \tag{38}$$

for some universal constant $D > 0$.

Proposition 5.14. *There exists $r_0 > 0$, so that for $r > r_0$, the following holds. Let $(z,v),(z',v') \in T^1\mathbf{H}$, such that $\mathbf{p}(z,v) = \mathbf{p}(z',v')$, and $\mathbf{d}(z,z') < 10$. Suppose that $\Delta_{\mathbf{p}(z,v)}[z,z'] = 0$. Let $p \in \partial\mathbf{H}$, and suppose $\Delta_p[\mathbf{z}_{\max}(\gamma_{(z,v)},p),z] \geq r$. Then*

$$\mathbf{d}(\mathbf{z}_{\max}(\gamma_{(z,v)},p),\mathbf{z}_{\max}(\gamma_{(z',v')},p)) \leq re^{-r}. \tag{39}$$

Moreover, we have

$$|\Delta_p[\mathbf{z}_{\max}(\gamma_{(z,v)},p),z] - \Delta_p[\mathbf{z}_{\max}(\gamma_{(z',v')},p),z']| < 2re^{-r}. \tag{40}$$

Proof. The rays $\gamma_{(z,v)}$ and $\gamma_{(z',v')}$ end at the same point at $\partial\mathbf{H}$. Since $\Delta_{\mathbf{p}(z,v)}[z,z'] = 0$, and since $\mathbf{d}(z,z') < 10$, for any $t > 0$, applying (38) we obtain

$$\mathbf{d}(\gamma_{(z,v)}(t),\gamma_{(z',v')}(t)) \leq 10De^{-t}. \tag{41}$$

Let $t_0 > 0$, so that $\gamma_{(z,v)}(t_0) = \mathbf{z}_{\max}(\gamma_{(z,v)},p)$. By Proposition 5.2 we have $t_0 > \Delta_p[\mathbf{z}_{\max}(\gamma_{(z,v)},p),z] \geq r$. Let $\zeta = \gamma_{(z,v)}(t_0-1)$, and $\zeta_1 = \gamma_{(z,v)}(t_0+1)$. Also, let $w = \gamma_{(z',v')}(t_0-1)$, and $w_1 = \gamma_{(z',v')}(t_0+1)$. By (41) we have $\mathbf{d}(\zeta,w),\mathbf{d}(\zeta_1,w_1) \leq 10e^{1-t_0}$, so it follows from the previous proposition that for r large enough, we have

$$\mathbf{d}(\mathbf{z}_{\max}(\gamma_{(z,v)},p),\mathbf{z}_{\max}(\gamma_{(z',v')},p)) \leq 10DCe^{1-t_0} < re^{-r}.$$

This shows that (39) holds.

We may assume that $\mathbf{p}(z,v) = \infty$. Let $\alpha : \mathbf{R} \to \mathbf{H}$, denote the naturally parametrised horocircle (with respect to ∞), so that the oriented angle between the vectors $\alpha'(0)$ and v, is $\frac{\pi}{2}$. Then for every s, we have

$$\Delta_\infty[z,\alpha(s)] = 0.$$

Moreover, there exists $s_0 \in \mathbf{R}$, so that $\alpha(s_0) = z'$. Assume $s_0 \geq 0$ (the other case is handled in the same way). The number s_0 depends only on the upper bound of the hyperbolic distance between z and z' which is bounded above by 10.

By $\alpha_r : \mathbf{R} \to \mathbf{H}$ denote the naturally parametrised horocircle (with respect to p), so that the oriented angle between the vectors $\alpha_r'(0)$ and v is positive. It follows from Proposition 5.2 that $|\Theta_p(z,v)| < \pi e^{-r}$. Therefore, the angle between the vectors $\alpha'(0)$ and $\alpha_r'(0)$ is at most πe^{-r}. We find that there exists a constant $K > 0$ that depend only on s_0 with the following properties. For $0 \leq s \leq (s_0+1)$ we have

$$|\Delta_\infty[z,\alpha_r(s)]| \leq Ke^{-r}, \tag{42}$$

Let $\gamma : \mathbf{R} \to \mathbf{H}$ denote the naturally parametrised geodesic that contains the geodesic ray $\gamma_{(z',v')}$ such that $\gamma(0) = z'$. For r large enough, we have that α_r intersects the geodesic γ. This is true because when $r \to \infty$ we have that the horoball at p that contains z converges on compact sets in $\mathbf{H}$ to the horoball at ∞ that contains z and on the other hand we have that α intersect γ orthogonally at z'. Let $z'' = \alpha_r \cap \gamma$. Let $0 \leq s'$, be such that $\alpha_r(s') = z''$. Then for r large enough, we have $0 \leq s' \leq (s_0+1)$. We have $\mathbf{d}(z',z'') = |\Delta_\infty[z',z'']| = |\Delta_\infty[z,z'']| = |\Delta_\infty[z,\alpha_r(s')]|$ so it follows from (42) that

$$\mathbf{d}(z',z'') \leq Ke^{-r}.$$

Note $\Delta_p[z,z''] = 0$. It follows from (39) that

$$|\Delta_p[\mathbf{z}_{\max}(\gamma_{(z,v)},p),z] - \Delta_p[\mathbf{z}_{\max}(\gamma,p),z'']| \leq \mathbf{d}(\mathbf{z}_{\max}(\gamma_{(z,v)},p),\mathbf{z}_{\max}(\gamma,p)) \leq re^{-r}.$$

Since $\mathbf{d}(z', z'') \leq Ke^{-r}$, we have for r large, that

$$|\Delta_p[\mathbf{z}_{\max}(\gamma_{(z,v)}, p), z] - \Delta_p[\mathbf{z}_{\max}(\gamma_{(z',v')}, p), z']| \leq re^{-r} + Ke^{-r} = 2re^{-r},$$

which proves (40).

□

The following proposition states that if two unit vectors in $T^1\mathbf{H}$ are nearby and related by the horocyclic flow then either they are absorbed by the same cusp or the first vector "just misses" being absorbed by some cusp or it just barely gets absorbed by some cusp.

Proposition 5.15. *Let $r > 4$ and let (z, v), $(z', v') \in T^1\mathbf{H}$ such that $(z, v) \in \mathcal{D}_r(G)$ and such that*

- $\mathbf{p}(z, v) = \mathbf{p}(z', v')$.
- $\mathbf{d}(z, z') \leq 10$ *and* $\Delta_{\mathbf{p}(z,v)}[z, z'] = 0$.

Assume that $\mathbf{a}_r(z, v) \neq \mathbf{a}_r(z', v')$. Then at least one of the following two conditions holds

1. *There exists a cusp $c \in \mathrm{Cusp}(G)$, such that*
$$r - r^2e^{-r} < \mathbf{h}_c(\mathbf{z}_{\max}(\gamma_{(z,v)}, c)) - \mathbf{h}_c(z) < r + r^2e^{-r},$$
and $\mathbf{z}_{\max}(\gamma_{(z,v)}, c) \in \mathcal{H}_c(1)$.
2. *There exists a cusp $c \in \mathrm{Cusp}(G)$, such that*
$$1 - r^2e^{-r} < \mathbf{h}_c(\mathbf{z}_{\max}(\gamma_{(z,v)}, c)) < 1 + r^2e^{-r},$$
and $\mathbf{z}_{\max}(\gamma_{(z,v)}, c) = \gamma_{(z,v)}(t)$, for some $r - 1 < t < r^2 + 1$.

Proof. Set $\mathbf{a}_r(z, v) = c \in \mathrm{Cusp}(G)$. We have $c \neq \mathbf{a}_r(z', v')$. First consider the case when the geodesic ray $\gamma_{(z',v')}$ does not intersect $\mathcal{H}_c(1)$. In this case we show that the condition (2) holds. It follows from (39) that

$$\mathbf{d}(\mathbf{z}_{\max}(\gamma_{(z,v)}, c), \mathbf{z}_{\max}(\gamma_{(z',v')}, c)) \leq re^{-r}.$$

Since $\mathbf{z}_{\max}(\gamma_{(z,v)}, c) \in \mathcal{H}_c(1)$ and since in this case $\mathbf{z}_{\max}(\gamma_{(z',v')}, c)$ does not belong to $\mathcal{H}_c(1)$, we conclude that $1 \leq \mathbf{h}_c(\mathbf{z}_{\max}(\gamma_{(z,v)}, c)) < 1 + r^2e^{-r}$. Let $t_0 > 0$ be such that $\mathbf{z}_{\max}(\gamma_{(z,v)}, c) = \gamma_{(z,v)}(t_0)$. In order to show that (2) holds it remains to prove that $r - 1 < t_0 < r^2 + 1$. Let $t_1 \geq 0$ so that $\gamma_{(z,v)}(t_1)$ is the first point of entry of the geodesic ray $\gamma_{(z,v)}$ into $\mathcal{H}_c(1)$. Then $\mathbf{t}_r(z, v) = t_1$ and by the assumption $(z, v) \in \mathcal{D}_r(G)$ we have $t_1 \leq r^2$. Since $\mathbf{h}_c(\gamma_{(z,v)}(t_0)) - \mathbf{h}_c(\gamma_{(z,v)}(t_1) \leq r^2e^{-r}$ and since by Proposition 5.2 we have $t_0 - t_1 < \mathbf{h}_c(\gamma_{(z,v)}(t_0)) - \mathbf{h}_c(\gamma_{(z,v)}(t_1) + \log 2$, we have $\mathbf{t}_r(z, v) \leq t_0 < \mathbf{t}_r(z, v) + 1 \leq r^2 + 1$. On the other hand, since $\mathbf{a}_r(z, v) = c$ we have $\mathbf{h}_c(\gamma_{(z,v)}(t_0)) - \mathbf{h}_c(z) \geq r$. Then from Proposition 5.2 we have $t_0 > r - 1$. This shows that the condition (2) holds.

From now on we assume that $\gamma_{(z',v')}$ enters the horoball $\mathcal{H}_c(1)$ but that $c \neq \mathbf{a}_r(z', v')$. Then there are two possible reasons why $c \neq \mathbf{a}_r(z', v')$. The first one is that $\Delta_c[\mathbf{z}_{\max}(\gamma_{(z',v')}, c), z'] = \mathbf{h}_c(\mathbf{z}_{\max}(\gamma_{(z',v')}, c)) - \mathbf{h}_c(z') < r$. Since $\Delta_c[\mathbf{z}_{\max}(\gamma_{(z,v)}, c), z] = \mathbf{h}_c(\mathbf{z}_{\max}(\gamma_{(z,v)}, c)) - \mathbf{h}_c(z) \geq r$, from (40) we find that $r \leq \mathbf{h}_c(\mathbf{z}_{\max}(\gamma_{(z,v)}, c)) - \mathbf{h}_c(z) < r + r^2e^{-r}$. So in this case the condition (1) holds because we already know that $\mathbf{z}_{\max}(\gamma_{(z,v)}, c) \in \mathcal{H}_c(1)$.

The second reason is that the geodesic ray $\gamma_{(z',v')}$ gets absorbed before entering the horoball $\mathcal{H}_c(1)$. Set $\mathbf{a}_r(z', v') = c' \in \mathrm{Cusp}(G)$. If $t_0 \geq 0$ is such that $\mathbf{z}_{\max}(\gamma_{(z,v)}, c') = \gamma_{(z,v)}(t_0)$ then $t_0 < \mathbf{t}_r(z, v) \leq r^2$. Assume that $\mathbf{z}_{\max}(\gamma_{(z,v)}, c')$

does not belong to $\mathcal{H}_{c'}(1)$. Then we show that the condition (2) holds (with respect to the cusp c'). Since $\mathbf{z}_{\max}(\gamma_{(z',v')},c') \in \mathcal{H}_{c'}(1)$ by (39) we have that $\mathbf{z}_{\max}(\gamma_{(z,v)},c') \in \mathcal{H}_{c'}(1-r^2e^{-r})$. Since $\mathbf{h}_{c'}(\mathbf{z}_{\max}(\gamma_{(z',v')},c')) - \mathbf{h}_{c'}(z') \geq r$, from (40) we get that

$$\mathbf{h}_{c'}(\gamma_{(z,v)}(t_0)) - \mathbf{h}_{c'}(z) > \mathbf{h}_{c'}(\mathbf{z}_{\max}(\gamma_{(z',v')},c')) - \mathbf{h}_{c'}(z') - r^2e^{-r} > r - r^2e^{-r}.$$

This yields that $r-1 < t_0$. We have already seen that $t_0 < \mathbf{t}_r(z,v) \leq r^2$. This proves the statement.

It remains to examine the case $\mathbf{z}_{\max}(\gamma_{(z,v)},c') \in \mathcal{H}_{c'}(1)$. We show that in this case (2) holds. Again let $t_0 \geq 0$ such that $\mathbf{z}_{\max}(\gamma_{(z,v)},c') = \gamma_{(z,v)}(t_0)$. It follows from (40) that

$$\mathbf{h}_{c'}(\gamma_{(z,v)}(t_0)) - \mathbf{h}_{c'}(z) > \mathbf{h}_{c'}(\mathbf{z}_{\max}(\gamma_{(z',v')},c')) - \mathbf{h}_{c'}(z') - r^2e^{-r} > r - r^2e^{-r}.$$

On the other hand, we have that $\mathbf{h}_{c'}(\gamma_{(z,v)}(t_0)) - \mathbf{h}_{c'}(z) < r$ because otherwise we would have that $\mathbf{a}_r(z,v) = c'$. The last two estimates put together give us that $r - r^2e^{-r} < \mathbf{h}_{c'}(\gamma_{(z,v)}(t_0)) - \mathbf{h}_{c'}(z) < r$. This proves the proposition.

□

The following proposition replaces the hypotheses $\mathbf{d}(z,z') \leq 10$ and $\Delta_{\mathbf{p}(z,v)}[z,z'] = 0$ of Proposition 5.15 with the condition that z and z' are in $\mathbf{Th}_G(2r-\log 3)$.

Proposition 5.16. *Let $r > 4$ and $c \in \operatorname{Cusp}(G)$. Let (z,v), $(z',v') \in T^1\mathcal{H}_c(2r-\log 3)$ where $(z,v) \in \mathcal{D}_r(G)$ such that $\mathbf{p}(z,v) = \mathbf{p}(z',v')$ and such that $\mathbf{a}_r(z,v) \neq c \neq \mathbf{a}_r(z',v')$. If $\mathbf{a}_r(z,v) \neq \mathbf{a}_r(z',v')$ then the following holds*

- *There exists a cusp $c \in \operatorname{Cusp}(G)$, such that*
$$1 - r^2e^{-r} < \mathbf{h}_c(\mathbf{z}_{\max}(\gamma_{(z,v)},c)) < 1 + r^2e^{-r},$$
and $\mathbf{z}_{\max}(\gamma_{(z,v)},c) = \gamma_{(z,v)}(t)$ for some $r-1 < t < r^2+1$.

Proof. Since $\mathbf{a}_r(z,v) \neq c \neq \mathbf{a}_r(z',v')$ we have that $\mathbf{p}(z,v) = \mathbf{p}(z',v') \neq c$. Since $2r - \log 3 > r+2$ (because $r > 4$), it follows from Proposition 5.6 that $\gamma_{(z,v)}$ gets absorbed by the first 1-horoball it hits after leaving $\mathcal{H}_c(1)$. The same is true for the ray $\gamma_{(z',v')}$.

Let η be the horocircle at $\mathbf{p}(z,v)$ that is tangent to $\mathcal{H}_c(2r-\log 3)$. Let γ_j, $j=1,2$, be the two geodesics that start at $\mathbf{p}(z,v)$ and that are tangent to $\mathcal{H}_c(2r-\log 3)$. Let η_1 be the subsegment of η that is bounded by the points $\gamma_j \cap \eta$ (since γ_j starts at the same point on $\mathbf{R}$ where η touches $\mathbf{R}$, there exists a unique intersection point $\gamma_j \cap \eta$ in $\mathbf{H}$). We have that the hyperbolic length of η_1 is equal to 1 and that η_1 is contained in $\mathcal{H}_c(r+2)$. Let z_1 and z_1' be the points of intersection between η_1 and the geodesic rays $\gamma_{(z,v)}$ and $\gamma_{(z',v')}$ respectively. Observe that $\Delta_{\mathbf{p}(z,v)}[z_1,z_1'] = 0$ and $\mathbf{d}(z_1,z_1') < 1$. Also $z_1, z_1' \in \mathcal{H}_c(r+2)$.

Let $(z_1,v_1) \in T^1\mathbf{H}$ be such that the ray $\gamma_{(z_1,v_1)}$ is contained in the ray $\gamma_{(z,v)}$. Similarly let $(z_1',v_1') \in T^1\mathbf{H}$ be such that the ray $\gamma_{(z_1',v_1')}$ is contained in the ray $\gamma_{(z'v')}$. Since $z_1, z_1' \in \mathcal{H}_c(r+2)$ from Proposition 5.6 we have that $\mathbf{a}_r(z_1,v_1) = \mathbf{a}_r(z,v) = c_1$ and $\mathbf{a}_r(z_1',v_1') = \mathbf{a}_r(z',v')$. Also $r+2 \leq \mathbf{t}_r(z_1,v_1) < \mathbf{t}_r(z,v) \leq r^2$. We now apply the previous proposition to (z_1,v_1) and (z_1',v_1'). Since $r+2 \leq \mathbf{t}_r(z_1,v_1)$ we see that the condition (1) from the previous proposition can not hold so we have that the condition (2) holds for (z_1,v_1) and hence for (z,v). This proves the proposition.

□

We can now define three "error sets " and show that any point of the "bad set" H_r belongs to one of these three sets.

Definition 5.11. *Let $r > 2$. Define the sets $\mathcal{E}_r^i(G) \subset \mathcal{D}_r(G)$, $i = 1, 2, 3$, as follows.*
We say that $(z, v) \in \mathcal{E}_r^1(G)$ if

- *We have $(z, v) \in \mathcal{D}_r(G)$.*
- *We have $\mathbf{h}(z) \geq 2r - \log 3$.*
- *Let $c \in \mathrm{Cusp}(G)$ such that $z \in \mathcal{H}_c(2r - \log 3)$. Then c is an endpoint of $\mathbf{a}_r^1(z, v)$.*

We say that $(z, v) \in \mathcal{E}_r^2(G)$ if

- *There exists a cusp $c \in \mathrm{Cusp}(G)$, such that $r - r^2 e^{-r} < \mathbf{h}_c(\mathbf{z}_{\max}(\gamma_{(z,v)}, c)) - \mathbf{h}_c(z) < r + r^2 e^{-r}$ and $\mathbf{z}_{\max}(\gamma_{(z,v)}, c) \in \mathcal{H}_c(1)$.*
- *We have $(z, v) \in \mathcal{D}_r(G)$ and (z, v) does not belong to $\mathcal{E}_r^1(G)$.*

We say that $(z, v) \in \mathcal{E}_r^3(G)$ if

- *There exists a cusp $c \in \mathrm{Cusp}(G)$, so that $1 - r^2 e^{-r} < \mathbf{h}_c(\mathbf{z}_{\max}(\gamma_{(z,v)}, c)) < 1 + r^2 e^{-r}$ and so that $\mathbf{z}_{\max}(\gamma_{(z,v)}, c) = \gamma_{(z,v)}(t)$ for some $r - 1 < t < r^2 + 1$.*
- *We have $(z, v) \in \mathcal{D}_r(G)$ and (z, v) does not belong to $E_r^1(G)$.*

Note that the set $\mathcal{E}_r^i(G)$, $i = 1, 2, 3$, is invariant under the action of G and the corresponding quotient is denoted by $\mathcal{E}_r^i(S)$.

Proposition 5.17. *Suppose that $(z, v) \in H_r$. Then $(z, v) \in \mathcal{E}_r^i(G)$, for some $i = 1, 2, 3$, or $(z, \omega v) \in \mathcal{E}_r^i(G)$ for some $i = 2, 3$.*

Proof. We refer the reader to the flow chart for the logic of the proof below. Assume that $(z, v) \in \big(T^1\mathbf{H} \setminus (\mathcal{A}_r(G) \cup \mathcal{B}_r(G))\big)$ and that $(z, v) \in \mathcal{D}_r(G)$.

If $(z, v) \in T^1\mathbf{Th}_G(2r) \cap \mathcal{R}\big(T^1\mathbf{Th}_G(2r)\big)$ then by the definition of $\mathcal{A}_r(G)$ we find that either $\mathbf{a}_r(z, v) \neq \mathbf{a}_r(\omega(\mathcal{R}(z, v)))$ or $\mathbf{a}_r(z, \omega v) \neq \mathbf{a}_r(\mathcal{R}(z, v))$. Suppose that $\mathbf{a}_r(z, v) \neq \mathbf{a}_r(\omega(\mathcal{R}(z, v)))$ (the other case is handled in the same way), and set $(z', v') = \omega(\mathcal{R}(z, v))$. Then $\mathbf{p}(z, v) = \mathbf{p}(z', v')$ and $\Delta_{\mathbf{p}(z,v)}[z, z'] = 0$. Also, $\mathbf{d}(z, z') = \log 3 < 10$. Then by Proposition 5.15 we have that (z, v) belongs to one of the sets $\mathcal{E}_r^j(G)$, $j = 2, 3$.

If $(z, v) \in T^1\mathbf{Thin}_G(2r) \cup \mathcal{R}\big(T^1\mathbf{Thin}_G(2r)\big)$ then by Proposition 5.10 we have $z \in \mathcal{H}_c(2r - \log 3)$ for some $c \in \mathrm{Cusp}(G)$. Suppose that (z, v) does not belong to $\mathcal{B}_r(G)$. There are two reasons why this can happen. The first one is that (z, v) does not belong to the set $\widetilde{Q}_c = Q_c \cap \mathcal{R}(Q_c)$ defined in Definition 5.6. Set $(z', v') = \mathcal{R}(z, v)$. Then $\mathbf{a}_r^1(z, v)$ or $\mathbf{a}_r^1(z', v')$ has c as its endpoint. If $\mathbf{a}_r^1(z, v)$ has c as its endpoint then $(z, v) \in \mathcal{E}_r^1(G)$. If $\mathbf{a}_r(z', \omega v') = c$ then by Proposition 5.15 we have that (z, v) belongs to one of the sets $\mathcal{E}_r^j(G)$, $j = 2, 3$. If $\mathbf{a}_r(z', v') = c$ then $(z, \omega v)$ belongs to one of the sets $\mathcal{E}_r^j(G)$, $j = 2, 3$.

Assume $(z, v) \in \widetilde{Q}_c$. If (z, v) does not belong to $\mathcal{B}_r(G)$ then there exists $(z', v') \in T^1\mathcal{H}_c(2r) \cup \mathcal{R}\big(T^1\mathcal{H}_c(2r)\big)$ such that

- $\mathbf{p}^1(z, v) = \mathbf{p}^1(z', v')$.
- $\mathbf{a}_r^1(z', v')$ does not have c as its endpoint.
- $\mathbf{a}_r^1(z, v) \neq \mathbf{a}_r^1(z', v')$.

Then $z, z' \in \mathcal{H}_c(2r - \log 3)$. The condition $\mathbf{a}_r^1(z, v) \neq \mathbf{a}_r^1(z', v')$ means that $\mathbf{a}_r(z, v) \neq \mathbf{a}_r(z', v')$ or $\mathbf{a}_r(z, \omega v) \neq \mathbf{a}_r(z', \omega v')$. Applying the previous proposition to (z, v) and

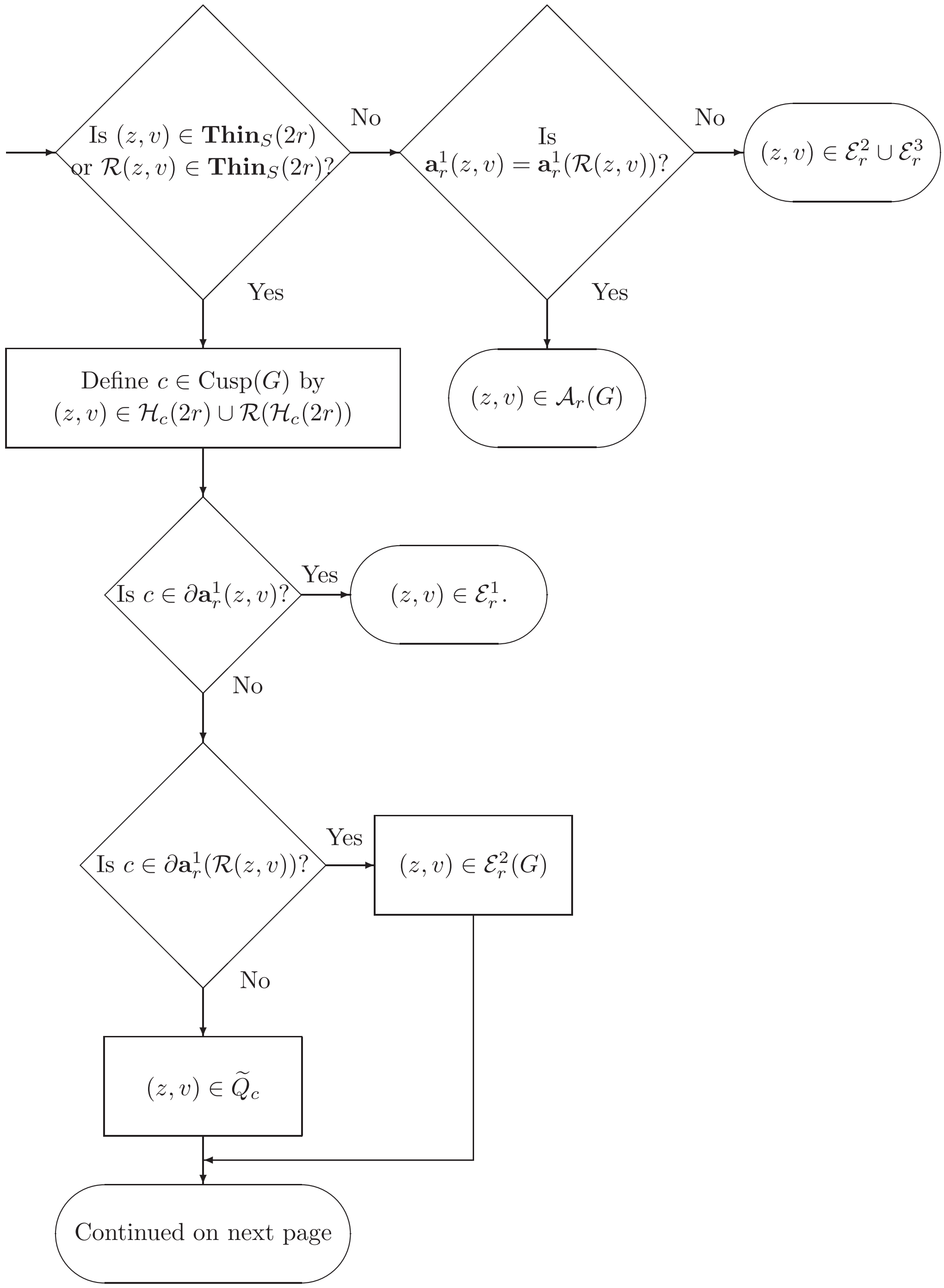

FIGURE 5. Flow chart part 1

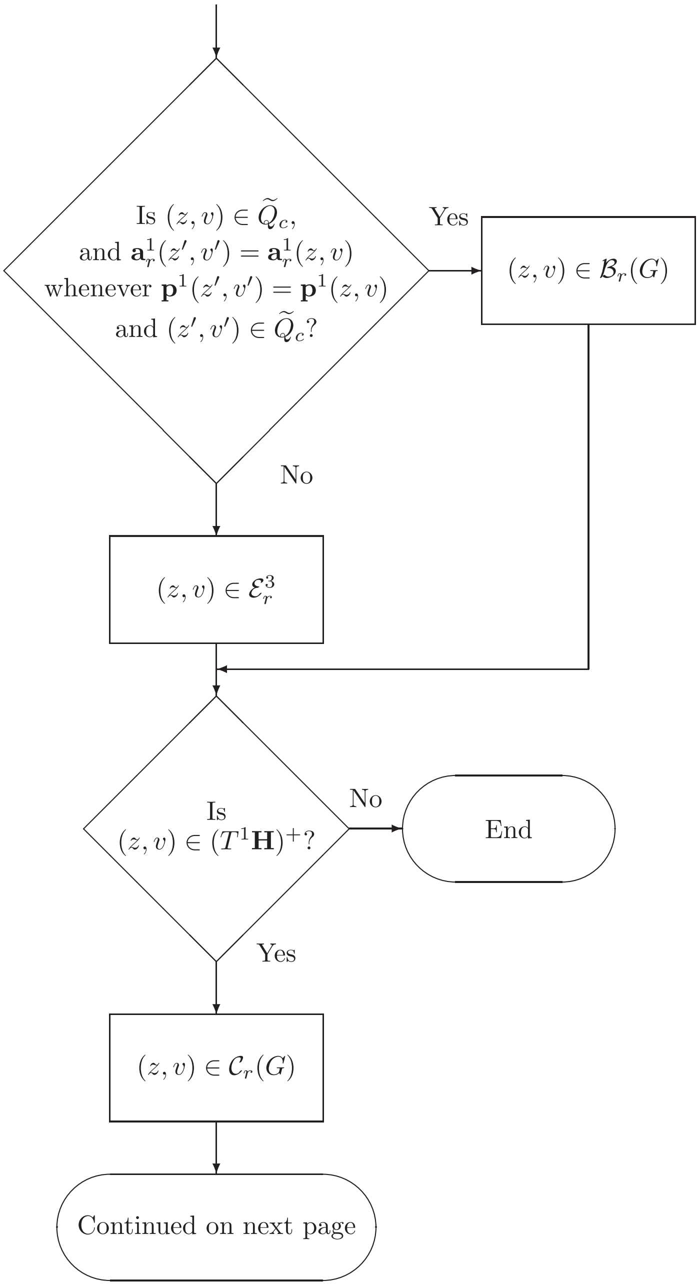

FIGURE 5. Flow chart part 2

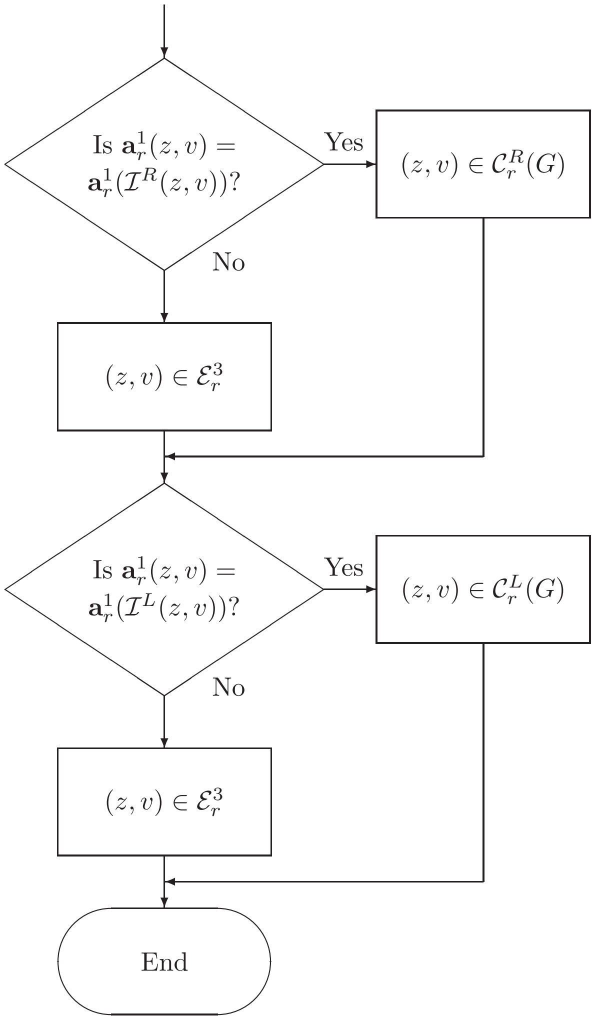

FIGURE 5. Flow chart part 3

(z',v') (or $(z,\omega v)$ and $(z',\omega v')$) we conclude that (z,v) or $(z,\omega v)$ belongs to the set $\mathcal{E}^3_r(G)$.

Assume that $(z,v) \in (\mathcal{C}_r(G) \setminus \mathcal{C}^L_r(G))$ and that $(z,v) \in \mathcal{D}_r(G)$. Then $z \in \mathcal{H}_c(2r)$ for some cusp $c \in \operatorname{Cusp}(G)$ and c is not an endpoint of $\mathbf{a}^1_r(z,v)$ (by the definition of $\mathcal{C}_r(G)$). Then $\mathbf{p}(z,v) \neq c \neq \mathbf{p}(z,\omega v)$. Set $(z',v') = \mathcal{I}^L(z,v)$ (note that $\mathcal{I}^L(z,v)$ is well defined since $\mathbf{p}(z,v) \neq c \neq \mathbf{p}(z,\omega v)$). Then $\mathbf{a}^1_r(z',v')$ since by Proposition 5.9 $(z',v') \in \mathcal{C}_r(G)$. Suppose that $\mathbf{a}_r(z,v) \neq \mathbf{a}_r(z',\omega v')$. Then by the previous proposition we have that $(z,v) \in \mathcal{E}^3_r(G)$. The case $(z,v) \in (\mathcal{C}_r(G) \setminus \mathcal{C}^R_r(G))$ is treated in the same way. $\square$

5.4. **The proof of Theorem 4.1.** Recall the statement of Theorem 4.1.

Theorem 5.1. *There exist a constant $r_0(S) = r_0$ that depends only on S, so that for every $r > r_0$ there exists a finite measure $\mu(r) \in \mathcal{M}(\mathcal{T}(S))$ so that the total measure $|\mu(r)|$ satisfies the inequality*

$$\frac{\Lambda(T^1S)}{2} < |\mu(r)| < \frac{3\Lambda(T^1S)}{2},$$

and with the following properties. There exist measures $\alpha(r), \alpha_1(r), \beta(r) \in \mathcal{M}(\mathbf{N}^1\Gamma(S))$ so that the measure $\widehat{\partial}\mu(r) \in \mathcal{M}(\mathbf{N}^1\Gamma(S))$ can be written as $\widehat{\partial}\mu(r) = \alpha(r) + \alpha_1(r) + \beta(r)$ and the following holds

1. *Let $\widehat{\mu}(r)$ denote the restriction of the measure $\mu(r)$ to the set of triangles $T \in \mathcal{T}(S)$ for which $\mathbf{h}(T) \geq r^2$. Then*
$$\int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1) d\widehat{\partial}\widehat{\mu}(r) \leq P(r)e^{-r}.$$
2. *The measure $\alpha(r)$ is re^{-r}-symmetric, and the measure $\alpha_1(r)$ is Q-symmetric, for any $Q > 200$.*
3. *We have*
$$\int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1) d\beta(r) \leq P(r)e^{-r}.$$
4. *We have*
$$\int\limits_{\mathbf{N}^1\Gamma(S)} d\alpha_1(r) \leq e^{-r}.$$

Proof. Let $\vartheta_r : T^1S \to \mathbf{R}$ be defined as follows. If (z, v) does not belong to $\mathcal{C}_r(S)$ then $\vartheta_r(z, v) = 0$. For $(z, v) \in \mathcal{C}_r(S)$ we have

$$\vartheta(z, v) = \varphi_r(\mathcal{R}(z, v)) - \varphi_r(z, v).$$

If $(z, v) \in \mathcal{C}_r(G)$ we have that $z \in \mathcal{H}_{\mathbf{b}_r(z,v)}(2r)$. Therefore, $\mathbf{h}(z) = \mathbf{h}_{\mathbf{b}_r(z,v)}(z)$. It follows from the definition of φ_r that $\varphi_r(z)$ depends only on that $\mathbf{h}_{\mathbf{b}_r(z,v)}(z)$ and is decreasing in $\mathbf{h}_{\mathbf{b}_r(z,v)}(z)$. Since $(z, v) \in \mathcal{C}_r(S) \subset (T^1S)^+$ we have that ϑ_r is a non-negative function.

We define the measure

$$\mu(r) = (\mathbf{a}_r^2)_*(\varphi_r \, d\Lambda) + 3(\mathbf{b}_r^2)_*(\vartheta_r \, d\Lambda). \tag{43}$$

Since $\varphi_r \leq 1$ on T^1S and since $\varphi_r(z, v) = 1$ for every $z \in \mathbf{Th}_S(r)$ we conclude that the total measure of $\varphi_r \, d\Lambda$ is approaching $\Lambda(T^1S)$ when $r \to \infty$. Since $\vartheta_r \leq 2$ on T^1S and since $\vartheta_r(z) = 0$ for every $(z, v) \in T^1\mathbf{Th}_S(r)$, we have that the total measure of $\vartheta_r \, d\Lambda$ tends to zero when $r \to \infty$. So for r large enough we have that the total measure $|\mu(r)|$ satisfies the inequality

$$\frac{\Lambda(T^1S)}{2} < |\mu(r)| < \frac{3\Lambda(T^1S)}{2}.$$

Let $\widehat{\mu}(r)$ be the restriction of the measure $\mu(r)$ to the set of triangles $T \in \mathcal{T}(S)$ for which $\mathbf{h}(T) \geq r^2$. Let $(z, v) \in T^1S$ be such that the triangle $\mathbf{a}_r^2(z, v)$ belongs to the support of $\widehat{\mu}(r)$. Then at least one of the inequalities $\mathbf{t}_r(z, \omega^j v) > r^2$, $j = 0, 1, 2$, is satisfied, so $(z,v) \in T^1S \setminus \mathcal{D}_r(S)$. From Proposition 5.12 and Lemma 5.2, we have that for r large enough, the following holds,

$$\int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*,z)+1)d\widehat{\partial}\widehat{\mu}(r) < 5 \int\limits_{T^1S\setminus\mathcal{D}_r(S)} (\mathcal{K}_r(z,v)+\mathcal{K}_r(z,\omega v)+\mathcal{K}_r(z,\omega^2 v)+1)\varphi_r(z)\,d\Lambda$$
$$< P(r)e^{-r}.$$

Next, we define the measures $\alpha(r)$, $\alpha_1(r)$, and $\beta(r)$. Let $\varphi_r^A(z,v) = \varphi_r(z)$ if $(z,v) \in \mathcal{A}_r(S)$, and 0 otherwise. Let $\varphi_r^B(z,v) = \varphi_r(z)$ if $(z,v) \in \mathcal{B}_r(S)$, and 0 otherwise. Let $\vartheta_r^B(z,v) = \vartheta_r(z,v)$ if $(z,v) \in \mathcal{B}_r(S)$, and 0 otherwise. Let $\vartheta_r^1(z,v) = \vartheta_r(z,v)$ if $(z,v) \in \mathcal{C}_r^L(S)$, and 0 otherwise, and let $\vartheta_r^2(z,v) = \vartheta_r(z,v)$ if $(z,v) \in \mathcal{C}_r^R(S)$, and 0 otherwise.

Set

$$\alpha'(r) = 3(\mathbf{f}_r)_*(\varphi_r^A\,d\Lambda),\ \ \alpha_1(r) = 3(\mathbf{f}_r)_*(\varphi_r^B\,d\Lambda) + 3(\bar{\mathbf{f}}_r)_*(\vartheta_r^B\,d\Lambda),$$

and

$$\alpha''(r) = 3(\bar{\mathbf{f}}_r^L)_*(\vartheta_r^1\,d\Lambda) + 3(\bar{\mathbf{f}}_r^R)_*(\vartheta_r^2\,d\Lambda).$$

Set $\alpha(r) = \alpha'(r) + \alpha''(r)$. In addition, let

$$\beta(r) = 3(\mathbf{f}_r)_*((\varphi_r - \varphi_r^A - \varphi_r^B)\,d\Lambda) + 3(\bar{\mathbf{f}}_r)_*(\vartheta_r - \vartheta_r^B)+$$
$$+3(\bar{\mathbf{f}}_r^L)_*((\vartheta_r - \vartheta_r^1)\,d\Lambda) + 3(\bar{\mathbf{f}}_r^R)_*((\vartheta_r - \vartheta_r^2)\,d\Lambda).$$

We have

$$\alpha(r)+\alpha_1(r)+\beta(r) = 3(\mathbf{f}_r)_*(\varphi_r\,d\Lambda)+3(\bar{\mathbf{f}}_r)_*(\vartheta_r\,d\Lambda)+3(\bar{\mathbf{f}}_r^L)_*(\vartheta_r\,d\Lambda)+3(\bar{\mathbf{f}}_r^R)_*(\vartheta_r\,d\Lambda),$$

so we conclude from (35) and (37) that $\widehat{\partial}\mu(r) = \alpha(r) + \alpha_1(r) + \beta(r)$. In order to prove Theorem 4.1, it remains to prove that the measures $\alpha(r)$, $\alpha_1(r)$, and $\beta(r)$ satisfy the corresponding properties.

We first show that $\alpha'(r)$ is re^{-r}-symmetric. Let $\gamma \in \Gamma(G)$ and let γ_1^* and γ_2^* denote the two orientations on γ. Let $Y_i(\gamma) \subset \mathcal{A}_r(G)$, $i = 1,2$, so that $(z,v) \in Y_i(\gamma)$ if $\mathbf{f}_r(z,v) \in N^1\gamma_i^*$. By the definition of the set $\mathcal{A}_r(G)$ we have that $Y_1(\gamma)$ and $Y_2(\gamma)$ are disjoint, and $\mathcal{R}(Y_1(\gamma)) = Y_2(\gamma)$. For $(z,v) \in Y_1(\gamma)$ set $\mathcal{R}(z,v) = (z_1,v_1)$. Combining Proposition 5.5 and Proposition 5.8, we have that

$$\mathbf{d}(\mathbf{f}_r(z,v),\mathbf{f}_r(z_1,v_1)) \leq \mathbf{d}(\mathbf{f}_r(z,v),\text{foot}_{\mathbf{a}_r^1(z,v)}(z)) + \mathbf{d}(\mathbf{f}_r(z_1,v_1),\text{foot}_{\mathbf{a}_r^1(z,v)}(z_1))$$
$$\leq Ce^{-r} + Ce^{-r} < re^{-r},$$

for r large enough. From the property (3) of Proposition 4.1, and since $\text{Jac}(\mathcal{R}) = 1$ we see that the measures $(\mathbf{f}_r)_*\nu_1$ and $(\mathbf{f}_r)_*\nu_2$ are re^{-r}-equivalent on γ, where ν_i is the restriction of the measure $3(\mathbf{f}_r)_*(\varphi_r\,d\Lambda)$ on $Y_i(\gamma)$. This shows that $\alpha'(r)$ is re^{-r}-symmetric.

Next, we show that $\alpha''(r)$ is e^{-r}-symmetric (and therefore this measure is re^{-r}-symmetric). Let $(z,v) \in \mathcal{C}_r(G)$. By Proposition 5.9 we have that $(z,v) \in \mathcal{C}_r^L(G)$ if and only if $\mathcal{I}^L(z,v) \in \mathcal{C}_r^R(G)$. Also, for almost every such (z,v) we have that $\mathbf{a}_r^1(z,v)$, $\mathbf{a}_r^1(\mathcal{I}^L(z,v))$, and $\mathbf{a}_r^1(\mathcal{I}^R(z,v))$ are well defined. We only need to consider such points. Then, by Proposition 5.11 we have that $\mathbf{d}(\bar{\mathbf{f}}_r^L(z,v),\bar{\mathbf{f}}_r^R(\mathcal{I}^L(z,v))) < e^{-r}$ for r large enough.

Let γ^* be the orientation on $\gamma = \mathbf{b}_r^1(z,v) = \mathbf{b}_r^1(\omega(\mathcal{I}^L(z,v)))$ so that the endpoint $\mathbf{a}_r(z,v)$ comes before the endpoint $\mathbf{b}_r(z,v)$ on γ^*. Let $Y_1(\gamma) \subset \mathcal{C}_r^L(G)$ so that $(z,v) \in Y_1$ if $\bar{\mathbf{f}}_r^L(z,v) \in N^1\gamma^*$. Set $Y_2(\gamma) = \mathcal{I}^L(Y_1(\gamma))$. We have $Y_2(\gamma) \subset \mathcal{C}_r^R(G)$

and for $(w,u) \in Y_2(\gamma)$ we have $\bar{\mathbf{f}}^R_r(w,u) \in N^1(-\gamma^*)$. Since $\mathrm{Jac}(\mathcal{I}^L) = 1$ and from the property (3) of Proposition 4.1, we see that the measures $(\bar{\mathbf{f}}^L_r)_*\nu_1$ and $(\bar{\mathbf{f}}^R_r)_*\nu_2$ are e^{-r}-equivalent on γ. Here ν_1 is the restriction of the measure $\vartheta_r \, d\Lambda$ on $Y_1(\gamma)$, and ν_2 is the restriction of the measure $\vartheta_r \, d\Lambda$ on $Y_2(\gamma)$.

Since both measures $\alpha'(r)$ and $\alpha''(r)$ are re^{-r}-symmetric, we see that $\alpha(r)$ is re^{-r}-symmetric.

From Proposition 5.10 we have that if $(z,v) \in \mathcal{B}_r(S)$ then $\mathbf{h}(z) \ge 2r - \log 3$. For r large enough, we have

$$\int\limits_{\mathbf{N}^1\Gamma(S)} d\alpha_1(r) \le 3 \int\limits_{\mathcal{B}_r(S)} (\varphi_r + \vartheta_r)\, d\Lambda \le 6\Lambda(T^1S \setminus T^1\mathbf{Th}_S(2r - \log 3)) \le e^{-r}.$$

The following lemma will be proved in the next section.

Lemma 5.4. *The measure $\alpha_1(r)$ is Q-symmetric, for any $Q > 200$.*

It remains to analyse $\beta(r)$. From the definition of $\beta(r)$ (and from (36)) we see that if either of the points $\mathbf{f}_r(z,v)$, $\bar{\mathbf{f}}_r(z,v)$, $\bar{\mathbf{f}}^L_r(z,v)$, or $\bar{\mathbf{f}}^R_r(z,v)$ belongs to the support of $\beta(r)$ we have that $(z,v) \in H_r$ (recall the definition of H_r from Lemma 5.3). Note that $\vartheta_r(z,v) \le \varphi_r(z,v)$ for every $(z,v) \in T^1S$. From Lemma 5.2, Lemma 5.3, and Proposition 5.12, we have

$$\int\limits_{\mathbf{N}^1\Gamma(S)} (\mathcal{K}(\gamma^*, z) + 1) d\beta(r) \le 5 \int\limits_{H_r} (\mathcal{K}_r(z,v) + \mathcal{K}_r(z, \omega v) + 1)\varphi_r(z)\, d\Lambda +$$

$$+5 \int\limits_{T^1S \setminus \mathcal{D}_r(S)} (\mathcal{K}_r(z,v) + \mathcal{K}_r(z,\omega v) + 1)\varphi_r(z)\, d\Lambda \le P(r)e^{-r},$$

for r large enough. This completes the proof of Theorem 4.1. □

6. The proof of Lemma 5.4

6.1. **Preliminary propositions.** We have the following preliminary propositions.

Proposition 6.1. *Let $(z,v) \in T^1\mathbf{H}$ and set $(z_1,v_1) = \mathcal{R}(z,v)$. Let $p \in \partial\mathbf{H}$. Set $w = \mathrm{foot}_\gamma(z) = \mathrm{foot}_\gamma(z_1)$ and denote by u a unit vector at w that is tangent to γ. Then*

$$|\Delta_p[z, z_1]| < 2|\Theta_p(w,u)| \sinh(\log \sqrt{3}).$$

Proof. Recall that $\mathbf{d}(w,z) = \mathbf{d}(w,z_1) = \log\sqrt{3}$. Let $f \in \mathbf{Mob}(\mathbf{H})$ be the rotation centred at w, and so that $\Theta_p(f(w,u)) = 0$. Then f is a rotation for the angle $|\Theta_p(w,u)|$. Let $(z',v') = f(z,v)$ and $(z'_1,v'_1) = f(z_1,v_1) = \mathcal{R}(z',v')$ (here we use that $\mathcal{R}$ commutes with f). Then $\Delta_p[z',z'_1] = 0$.

Recall $\mathbf{d}(w,z) = \frac{1}{2}\log 3$. Using the formula that says that the hyperbolic circumference of the circle of radius s is $2\pi\sinh(s)$ we conclude that

$$\mathbf{d}(z,z'), \mathbf{d}(z_1,z'_1) < |\Theta_p(w,u)| \sinh(\log\sqrt{3}).$$

Since $\Delta_p[z',z'_1] = 0$ and from the previous inequality we get

$$|\Delta_p[z,z_1]| < |\Delta_p[z',z'_1]| + \mathbf{d}(z,z') + \mathbf{d}(z_1,z'_1) < 2|\Theta_p(w,u)|\sinh(\log\sqrt{3}).$$

□

Let η be a geodesic in $\mathbf{H}$ and let $\phi : \mathbf{R} \to \eta$ be the natural parametrisation of η so that $\phi(0) = \mathbf{z}_{\max}(\eta, \infty)$, and so that ∞ is to the left of η^*, where η^* is the push-forward of the orientation on $\mathbf{R}$ by the map ϕ. Let $(z_0, v_0) \in \mathbf{H}$ be the unique point so that $\mathbf{p}^1(z_0, v_0) = \eta$, and so that $\Theta_\infty(z_0, \omega^2 v_0) = 0$. Set $\mathcal{S}_t(z_0, v_0) = (z_t, v_t)$. Let $(z_0', v_0') = \mathcal{R}(z_0, v_0)$, and set $\mathcal{S}_{(-t)}(z_0', v_0') = (z_t', v_t')$. Then $\text{foot}_\eta(z_t') = \text{foot}_\eta(z_t) = \phi(t) \in \eta$.

Definition 6.1. *Let* $q(t) = \mathbf{h}_\infty(z_t)$*, and* $p(t) = \mathbf{h}_\infty(z_t')$.

Proposition 6.2. *We have* $0 < q(t) - p(t) \leq 2\pi \sinh(\log \sqrt{3}) e^{-|t|}$*, and* $|q(t) - (q(0) - |t|)| < \log 6$.

Proof. Note that $q(t) - p(t) = \mathbf{h}_c(z_t) - \mathbf{h}_c(z_t') = \Delta_c[z_t, z_t']$. Let $w_t = \text{foot}_\eta(z_t) = \text{foot}_\eta(z_t')$, and let u_t be the unit vector so that $0 < |\Theta_c(w_t, u_t)| \leq \frac{\pi}{2}$. Such vector exists because c is not an endpoint of η (when $t = 0$ there are two such vectors and we choose either one). Also, $\mathbf{d}(w_0, w_t) = |t|$. From Proposition 6.1 we have

$$q(t) - p(t) < 2|\Theta_c(w_t, u_t)| \sinh(\log \sqrt{3}).$$

The identity (29) yields

$$|t| = \log\left(\csc(\Theta_c(w_t, u_t)) + \cot(\Theta_c(w_t, u_t))\right),$$

and therefore, we have

$$\frac{\sin(\Theta_c(w_t, u_t))}{2} \leq e^{-|t|} \leq \sin(\Theta_c(w_t, u_t)).$$

Since for $0 \leq \theta \leq \frac{\pi}{2}$, we have $\theta \leq \frac{\pi}{2} \sin\theta$, we find $|\Theta_c(w_t, u_t)| \leq \pi e^{-|t|}$. This shows

$$q(t) - p(t) < \pi e^{-|t|} 2 \sinh(\log \sqrt{3}) = 2\pi \sinh(\log \sqrt{3}) e^{-|t|},$$

which proves the first inequality.

It follows from Proposition 5.2 that $\mathbf{h}_c(w_0) - \mathbf{h}_c(w_t) > |t| - \log 2$. Since $\mathbf{d}(w_t, z_t) = \log \sqrt{3}$, we have $q(0) - q(t) > |t| - \log 2 - \log 3$, which proves $|q(t) - (q(0) - |t|)| < \log 6$. $\square$

Proposition 6.3. *Let* $T \geq 0$ *and let*

$$\delta_T(\eta) = \int_{-T}^{T} \left(e^{\frac{2r - p(t)}{2}} - e^{\frac{2r - q(t)}{2}}\right) dt.$$

Then

$$\delta_T(\eta) \leq 24(e^{\frac{\pi}{\sqrt{3}}} - 1) e^{\frac{2r - q(0)}{2}}$$

Proof. We compute

$$\delta_T(\eta) \leq \int_{-\infty}^{\infty} \left(e^{\frac{2r - p(t)}{2}} - e^{\frac{2r - q(t)}{2}}\right) dt = 2e^{\frac{2r - q(0)}{2}} \int_0^\infty e^{\frac{q(0) - q(t)}{2}} \left(e^{\frac{q(t) - p(t)}{2}} - 1\right) dt.$$

We apply Proposition 6.2 to the right hand side of the above inequality and since $e^{kx} - 1 \leq (e^k - 1)x$ for $x \in [0, 1]$, we get

$$\delta_T(\eta) \leq 2e^{\frac{2r - q(0)}{2}} \int_0^\infty 6 e^{\frac{t}{2}} \left(e^{\frac{\pi e^{-t}}{\sqrt{3}}} - 1\right) dt \leq 12\left(e^{\frac{\pi}{\sqrt{3}}} - 1\right) e^{\frac{2r - q(0)}{2}} \int_0^\infty e^{\frac{t}{2}} e^{-t}\, dt.$$

The identity

$$\int_0^\infty e^{-\frac{t}{2}}\, dt = 2,$$

completes the proof of the proposition. □

Proposition 6.4. *For any $T \geq 0$ let $\mu^+_{\eta,T}$ and $\mu^-_{\eta,T}$ be the measures on $\mathbf{R}$ that are supported on $[-T,T]$ and given by*

$$\mu^+_{\eta,T} = e^{\frac{2r-q(t)}{2}}\, dt,$$

and

$$\mu^-_{\eta,T} = e^{\frac{2r-p(t)}{2}}\, dt,$$

Let $\delta_{\eta,T}$ be the measure on $\mathbf{R}$ that is supported at the point 0 and such that

$$\delta_{\eta,T}(\mathbf{R}) = \delta_{\eta,T}(\{0\}) = \mu^-_{\eta,T}(\mathbf{R}) - \mu^+_{\eta,T}(\mathbf{R}).$$

Then the measures $(\mu^+_{\eta,T} + \delta_{\eta,T})$ and $\mu^-_{\eta,T}$ are Q-equivalent for $Q = 24\big(e^{\frac{\pi}{\sqrt{3}}} - 1\big)$.

Proof. We write μ^+, μ^- and δ for $\mu^+_{\eta,T}$, $\mu^-_{\eta,T}$ and $\delta_{\eta,T}$ respectively. To prove that the corresponding measures are Q-equivalent we use Proposition 4.1. Observe that for any $t \in \mathbf{R}$,

$$\mu^+(-\infty,t] \leq \mu^-(-\infty,t], \tag{44}$$

and

$$\delta(-\infty,t] \leq Qe^{\frac{2r-q(0)}{2}}. \tag{45}$$

Moreover when $-T \leq t \leq t+Q \leq T$ we have

$$\mu^-(t,t+Q) \geq \int_t^{t+Q} e^{\frac{2r-q(0)}{2}}\, ds = Qe^{\frac{2r-q(0)}{2}}. \tag{46}$$

Finally by the definition of δ we have

$$(\mu^+ + \delta)(\mathbf{R}) = \mu^-(\mathbf{R}). \tag{47}$$

We now show that for any $t \in \mathbf{R}$ we have

$$(\mu^+ + \delta)(-\infty,t] \leq \mu^-(-\infty,t+Q). \tag{48}$$

For $t \leq -T$ the left hand side of the above inequality is 0. For $-T \leq t < t+Q \leq T$ the inequality follows from (44), (45) and (46). For $t > T$ the inequality follows from (47). This shows (48).

By $s \to -s$ symmetry we have

$$\mu^+ + \delta)[-t,\infty) \leq \mu^-(-t-Q,\infty), \tag{49}$$

and replacing $-t$ with $t+Q$ we obtain

$$(\mu^+ + \delta)[t+Q,\infty) \leq \mu^-(t,\infty). \tag{50}$$

Thus by (47) we have

$$\mu^-(-\infty,t] \leq (\mu^+ + \delta)(-\infty,t+Q). \tag{51}$$

We have verified the hypothesis from Proposition 4.1 and this proves the proposition.

□

6.2. **The proof of Lemma 5.4.** Let $r > 2$ and let $\gamma \in \Gamma(G)$. We need to show that the restriction of the measure $\alpha_1(r)$ to $N^1\gamma$ is Q-symmetric. For r large, we have that the support of the measure $\alpha_1(r)$ is deep into the thin part of $N^1\gamma$. In fact, it follows from Proposition 5.10 that $\mathcal{B}_r(G) \subset T^1\mathbf{Thin}_G(2r - \log 3)$. Note that for $(z, v) \in \mathcal{B}_r(G)$ we have that $\mathbf{h}(\text{foot}_{\mathbf{p}^1(z,v)}(z) \geq 2r - \log 3$. Combining this with Proposition 5.5, we have that for r large enough, the restriction of the measure $\alpha_1(r)$ to $N^1\gamma$ is supported in $N^1\gamma \cap (T^1\mathbf{H} \setminus T^1\mathbf{Th}_G(2r - \log 3 - 1))$. Therefore, for each cusp $c \in \text{Cusp}(G)$ we consider the restriction of the measure $\alpha_1(r)$ to $N^1\gamma \cap T^1\mathcal{H}_c(2r - \log 3 - 1)$ which we denote by $\alpha_1(\gamma, c, r)$. To prove Lemma 5.4, it is enough to show that each $\alpha_1(\gamma, c, r)$ is Q-symmetric. There are only finitely many cusps $c \in \text{Cusp}(G)$, so that $\alpha_1(\gamma, c, r)$ is a non-zero measure, and the restriction of $\alpha_1(r)$ to $N^1\gamma$ is the finite sum of these measures $\alpha_1(\gamma, c, r)$. Note that if c is an endpoint of γ then $\alpha_1^c(r)$ is the zero measure.

Let $G = G_c$ (then $c = \infty$). We say that $(z, v) \in \mathcal{B}_r^*(\gamma, c)$, where $* \in \{+, -\}$, if the following holds

- We have $(z, v) \in \mathcal{B}_r(G)$.
- We have $(z, v) \in (T^1\mathbf{H})^*$.
- $\mathbf{f}_r(z, v) \in N^1(\gamma \cap \mathcal{H}_c(1))$.

We say that $(z_0, v_0) \in U^+ \subset \mathcal{B}_r^+(\gamma, c)$, if $\Theta_c(z_0, \omega^2 v_0) = 0$. Set $(z_0', v_0') = \mathcal{R}(z_0, v_0)$, and let $U^- = \mathcal{R}(U^+)$. By definition of $\mathcal{B}_r(G)$, we have $U^- \subset \mathcal{B}_r^-(\gamma, c)$. For $(z_0, v_0) \in \mathcal{U}^+$ and $t \in \mathbf{R}$ set $(z_t, v_t) = \mathcal{S}_t(z_0, v_0)$. Also $(z_t', v_t') = \mathcal{S}_{(-t)}(z_0', v_0')$. We have

Proposition 6.5. *Fix $(z_0, v_0) \in U^+$. Then there exists $T = T(z_0, v_0) \geq 0$ such that the set $t \in \mathbf{R} : (z_t, v_t) \in \mathcal{B}_r^+(\gamma, c)$ is a symmetric interval $(-T, T)$ or $[-T, T]$.*

Proof. Observe that $\mathbf{p}(z_0, v_0) = \mathbf{p}(z_t, v_t)$, for any $t \in \mathbf{R}$. The condition $(z_t, v_t) \in \mathcal{B}_r^+(\gamma, c)$ it equivalent with the following two conditions. The first one is that c is not an endpoint of $\mathbf{a}_r^1(z_t, v_t)$ or of $\mathbf{a}_r^1(z_t', v_t')$. The second one is that $(z_t, v_t) \in \big(\mathcal{H}_c(2r) \cup \mathcal{R}(\mathcal{H}_c(2r))\big)$. One directly verifies that each of these two conditions is satisfied on a symmetric interval. □

Proposition 6.6. *There exists $r_0 > 0$, so that for $r > r_0$, the following holds. Let $(z_0, v_0) \in U^+$. For every $t \in (-T(z_0, v_0), T(z_0, v_0))$, we have*

$$\mathbf{d}(\psi(t), \mathbf{f}_r(z_t, v_t)), \mathbf{d}(\psi(t), \mathbf{f}_r(z_t', v_t')) \leq re^{-r},$$

where $\psi : \mathbf{R} \to \gamma$ is the natural parametrisation of γ such that $\psi(0) = \mathbf{z}_{\max}(\gamma, c)$.

Proof. Assume that $G = G_c$ (then $c = \infty$). Since $(z_0, v_0) \in \mathcal{B}_r(G)$, and since (z_0, v_0) is above $\mathbf{p}^1(z_0, v_0)$, we have that $\mathbf{h}_\infty(z_0) \geq 2r > r + 2$. By Proposition 5.6 we have that $\mathcal{H}_{\mathbf{a}_r(z_0,v_0)}(1)$, is the first 1-horoball that the geodesic ray $\gamma_{(z_0,v_0)}$ enters after leaving $\mathcal{H}_\infty(1)$. Since with this normalisation the Euclidean diameter of any 1-horoball (except the one based at ∞) is at most 1, and since the ray $\gamma_{(z_0,v_0)}$ ends at $\mathbf{p}(z_0, v_0)$, we have $|\mathbf{a}_r(z_0, v_0) - \mathbf{p}(z_0, v_0)| < 1$. Similarly, $|\mathbf{a}_r(z_0, \omega v_0) - \mathbf{p}(z_0, \omega v_0)| < 1$. From $\mathbf{h}_\infty(z_0) \geq 2r$, we have that

$$|\mathbf{p}(z_0, v_0) - \mathbf{p}(z_0, \omega v_0)| \geq 2e^{-\mathbf{h}_\infty(z_0) - \log\sqrt{3}} > e^{2r}.$$

Let $f \in \mathbf{Mob}_\infty(\mathbf{H})$, be the unique Möbius transformation so that $f(\mathbf{p}^1(z_0, v_0)) = \gamma$ (note that $\text{foot}_\gamma(f(z_0)) = \psi(0)$). Then for $\zeta \in \mathbf{H}$, we have $f(\zeta) = l_1\zeta + l_2$, where

$l_1 > 0$, and $l_2 \in \mathbf{R}$. The coefficients l_1, l_2, are determined by the conditions $f(\mathbf{p}(z_0, v_0)) = \mathbf{a}_r(z_0, v_0)$, and $f(\mathbf{p}(z_0, \omega v_0)) = \mathbf{a}_r(z_0, \omega v_0)$. We have

$$|\log l_1| \leq \log(1 + 2e^{-2r}) < 2e^{-2r}$$

and $|l_2| < 1$. Since $\mathbf{h}_\infty(z_t), \mathbf{h}_\infty(z'_t) > 2r - \log 3$, for $t \in [-T(z_0, v_0), T(z_0, v_0)]$ we have that

$$\mathbf{d}(z_t, f(z_t)), \mathbf{d}(z'_t, f(z'_t)) \leq 2e^{-2r} + 3e^{-2r} = 5e^{-2r}.$$

Since $\text{foot}_\gamma(f(z_t)) = \psi(t)$ and since foot_γ does not increase the hyperbolic distance we obtain

$$\mathbf{d}(\text{foot}_\gamma(z_t), \psi(t)), \mathbf{d}(\text{foot}_\gamma(z'_t), \psi(t)) \leq 2e^{-2r} + 3e^{-2r} = 5e^{-2r}.$$

Together with Proposition 5.8 this completes the proof. □

Let $(z_0, v_0) \in U^+$. Define the mappings $f^{+/-}_{(z_0,v_0)} : (-T(z_0, v_0), T(z_0, v_0)) \to N^1\gamma$, by

$$f^{+}_{(z_0,v_0)}(t) = \mathbf{f}_r(z_t, v_t), \qquad f^{-}_{(z_0,v_0)}(t) = \mathbf{f}_r(z'_t, v'_t),$$

and let

$$\mu^{+/-}_{(z_0,v_0)} = (f^{+/-}_{(z_0,v_0)})_* (\mu^{+/-}_{\mathbf{p}^1(z_0,v_0),T(z_0,v_0)}),$$

where the measures $\mu^{+/-}_{\eta,t}$ were defined in Proposition 6.4. We have that $\mu^{+/-}_{(z_0,v_0)} \in \mathcal{M}(N^1\gamma)$. Note that $\mu^{+}_{(z_0,v_0)}$ is supported at the point $(\psi(t), \vec{n}(t))$, where $\vec{n}(t)$ is the unit normal vector at z that points towards the cusp c (so $|\Theta_c(\psi(t), \vec{n}(t))| < \frac{\pi}{2}$). The measure $\mu^{-}_{(z_0,v_0)}$ is supported at the point $(\psi(t), \vec{n}(t))$, where $\vec{n(t)}$ is the unit normal vector at z that points away from the cusp c (so $|\Theta_c(\psi(t), \vec{n}(t))| > \frac{\pi}{2}$).

Let $\delta_{(z_0,v_0)}$ be the atomic measure that on $N^1\gamma$ that is supported at the point $(\psi(0), \vec{n}_0)$, where $\Theta_c(\psi(0), \vec{n}_0) = 0$, and the mass of $\delta_{(z_0,v_0)}$ is $\delta_{T(z_0,v_0)}(\mathbf{p}^1(z_0, v_0))$ (recall the definition of $\delta_T(\eta)$ from Proposition 6.3).

Proposition 6.7. *For every* $(z_0, v_0) \in U^+$ *the measure* $(\mu^{+}_{(z_0,v_0)} + \delta_{(z_0,v_0)}) + \mu^{-}_{(z_0,v_0}$ *is* $Q + 1$*-symmetric.*

Proof. We can by abuse of notation write $\mu^{+/-}_{(z_0,v_0)}$ and $\delta_{(z_0,v_0)}$ as measures on γ rather than on $N^1\gamma$. Then we must show that $(\mu^{+}_{(z_0,v_0)} + \delta_{(z_0,v_0)})$ and $\mu^{-}_{(z_0,v_0}$ are $Q + 1$-equivalent.

We let $\eta = \mathbf{p}^1(z_0, v_0)$ and $T = T(z_0, v_0)$. By the previous proposition for $t \in (-T, T)$ we have $\mathbf{d}(f^{+/-}_{(z_0,v_0}(t), \psi(t)) \leq re^{-r} < \frac{1}{2}$, for r large enough. Therefore by Proposition 4.1 the measures $\mu^{+/-}_{(z_0,v_0)}$ and $\psi_* \mu^{+/-}_{\eta,T}$ are $\frac{1}{2}$-equivalent. Also $\delta_{(z_0,v_0)} = \psi_* \delta_{\eta,T}$. The proof now follows from Proposition 6.4 and Proposition 4.6. □

Let E be a Borel subset of U^+. For $t_1, t_2 \in \mathbf{R}$, $t_1 \leq t_2$, let

$$E(t_1, t_2) = \{\mathcal{S}_t(z, v) : (z, v) \in U^+, t_1 \leq t \leq t_2\}.$$

Since $\mathcal{S}_t$ is a measure preserving flow (one-parameter Abelian group) and since $\mathcal{S}_t(U^+) \cap U^+ = \emptyset$, there exists a non-negative constant $\sigma^+(E)$ that depends only on E, so that $\Lambda(E(t_1, t_2)) = \sigma^+(E)(t_2 - t_1)$. The constant $\sigma^+(E)$ is called the cross sectional area of the set E. This way we construct a positive Borel measure σ^+ on U^+. Similarly, we get the measure σ^- on U^-. If $E \subset U^+$, then $\mathcal{R}(E) \subset U^-$, and $\sigma^+(E) = \sigma^-(\mathcal{R}(E))$ (because $\text{Jac}(\mathcal{R}) = 1$, and $\mathcal{R} \circ \mathcal{S}_t = \mathcal{S}_{(-t)} \circ \mathcal{R}$). This shows that the restriction of the map $\mathcal{R}: U^+ \to U^-$, satisfies that $\mathcal{R}^*\sigma^- = \sigma^+$. If $\sigma^+(U^+) = 0$, then the measure $\alpha_1^c(r)$ is the zero measure.

It follows from the definition of $\alpha_1(r)$ and from (36) that

$$\alpha_1(\gamma, c, r) = 3 \int\limits_{U^+} \left((\mu^+_{(z_0,v_0)} + \delta_{(z_0,v_0)}) + \mu^-_{(z_0,v_0)} \right) d\sigma^+(z_0, v_0).$$

Then Lemma 5.4 follows from the previous proposition and Proposition 4.7. One sees from Proposition 6.4 that $Q + 1 \leq 200$.

7. Estimates on the combinatorial length

The following definition and propositions will be used throughout this section. Recall that for $(z, v) \in T^1\mathbf{H}$ the map $\gamma_{(z,v)} : [0, \infty) \to \mathbf{H}$ is the natural parametrisation of the geodesic ray that starts at z and that is tangent to the vector v at z that is $\gamma'_{(z,v)}(0) = v$. By $\gamma_{(z,v)}[0, t]$ we denote the corresponding geodesic segment.

Definition 7.1. *Let $(z, v) \in T^1\mathbf{H}$ and $t \in \mathbf{R}$. The geometric intersection number $\iota(\gamma_{(z,v)}(t), \tau(G))$ is defined as the number of (transverse) intersections between the geodesic segment $\gamma_{(z,v)}[0, t]$ and the edges from $\lambda(G)$. For $(z, v) \in T^1S$ the intersection number $\iota(\gamma_{(z,v)}(t), \tau(S))$ is defined in the same way.*

Next we establish the connection between the r-combinatorial length $\mathcal{K}_r(z, v)$ (Definition 5.7) and the intersection number $\iota(\gamma_{(z,v)}(\mathbf{t}_r(z, v)), \tau(S))$.

Proposition 7.1. *Let $(z, v) \in T^1S$ such that $0 < \mathbf{t}_r(z, v) < \infty$. Then*

$$\mathcal{K}_r(z, v) \leq \iota(\gamma_{(z,v)}(\mathbf{t}_r(z, v)), \tau(S)).$$

Proof. Let (z, v) also denotes a lift of (z, v) in $T^1\mathbf{H}$. Let γ be the geodesic ray that connects z and $\mathbf{a}_r(z, v) \in \operatorname{Cusp}(G)$. The assumption $\mathbf{t}_r(z, v) > 0$ implies that z does not belong to the horoball $\mathcal{H}_{\mathbf{a}_r(z,v)}(1)$ and that is why $\mathcal{K}_r(z, v) = \iota(\gamma, \tau(G))$. Let z_1 be the intersection point between γ and the horocircle $\partial\mathcal{H}_{\mathbf{a}_r(z,v)}(1)$ (since γ ends at $\mathbf{a}_r(z, v)$ there is exactly only one such intersection point z_1). Let γ_1 be the geodesic sub-segment of γ bounded by the points z and z_1. By definition the point $\gamma_{(z,v)}(\mathbf{t}_r(z, v))$ also belongs to the horocircle $\partial\mathcal{H}_{\mathbf{a}_r(z,v)}(1)$. Let $G = G_{\mathbf{a}_r(z,v)}$. Then the Euclidean distance between the points z_1 and $\gamma_{(z,v)}(\mathbf{t}_r(z, v))$ is at most 1 and the corresponding y-coordinate of both these points is equal to e.

Let α be an edge from $\lambda(G)$ such that γ intersects α. Then α does not end at ∞ (since γ ends at ∞ we see that γ can not (transversely) intersect any other geodesic that ends at ∞). Moreover we have that the segment γ_1 intersects α. But then α either intersects intersects the segment $\gamma_{(z,v)}[0, \mathbf{t}_r(z, v)]$ or the horocyclic segment between the points z_1 and $\gamma_{(z,v)}(\mathbf{t}_r(z, v))$. Every point on this horocyclic segment has the y-coordinate equal to e so we conclude that if α intersects this horocyclic segment it has to end at ∞. Therefore we find that α intersects the segment $\gamma_{(z,v)}[0, \mathbf{t}_r(z, v)]$. This proves the proposition. □

Proposition 7.2. *Let $(z, v) \in T^1S$. There exists a constant $K(S) > 0$ that depends only on S such that is $z \in \mathbf{Th}_S(1)$ then $\iota(\gamma_{(z,v)}(1), \tau(S)) < K(S)$. If $z \in (T^1S \setminus \mathbf{Th}_S(1))$ then $\iota(\gamma_{(z,v)}(1), \tau(S)) \leq N(S)e^{\mathbf{h}(z)}$. In particular for any $(z, v) \in T^1S$ we have $\iota(\gamma_{(z,v)}(1), \tau(S)) \leq \max\{K(S), N(S)e^{|\mathbf{h}(z)|}\}$.*

Proof. If $z \in \mathbf{Th}_S(1)$ then the geodesic segment remains in $\mathbf{Th}_S(2)$. Since the space of geodesic segments of the hyperbolic length 1 that remain in $\mathbf{Th}_S(2)$ is compact, we find that there exists $K(S) > 0$ so that every such segment intersects at most $K(S)$ edges from $\lambda(S)$.

Assume that z does not belong to $\mathbf{Th}_S(1)$ (this equivalent to saying that $\mathbf{h}(z) > 1$). Then $z \in \mathcal{H}_{c_i(S)}(1)$ for exactly one cusp $c_i(S) \in \operatorname{Cusp}(S)$. Let $G = G_{c_i(S)}$. Let $(z,v) \in T^1\mathbf{H}$ also denote a lift of (z,v) so that $z \in \mathcal{H}_\infty(1)$. Then the geodesic segment $\gamma_{(z,v)}[0,1]$ can intersect only edges from $\lambda(G)$ that end at ∞. This is why $\gamma_{(z,v)}[0,1]$ can intersect at most as many edges from $\lambda(G)$ as the horocyclic segment (horocyclic with respect to ∞) of the hyperbolic length 1 that begins at z. Therefore $\gamma_{(z,v)}[0,1]$ intersects at most $N(S)e^{\mathbf{h}_\infty(z)} = N(S)e^{\mathbf{h}(z)}$ edges from $\tau(G)$. This is easily seen by observing that $e^{\mathbf{h}_\infty(z)} \geq 1$ agrees with the y-coordinate of the point $z \in \mathbf{H}$. □

Proposition 7.3. *There exists a constant $D(S) > 0$ (that depends only on S) and $r_0 > 0$ so that for $r > r_0$ the following holds. Let $(z,v) \in T^1\mathbf{H}$. If $z \in \mathbf{Th}_G(10r)$ then*

$$\mathcal{K}_r(z,v) \leq D(S)\big(e^{11r} + K(S)(\mathbf{t}_r(z,v)+1)\big). \tag{52}$$

If $z \in \mathcal{H}_c(10r)$ for some $c \in \operatorname{Cusp}(G)$ then

$$\mathcal{K}_r(z,v) \leq D(S)\big(\chi_r(z,v)e^{\mathbf{h}(z)} + K(S)(\mathbf{t}_r(z,v)+1)\big). \tag{53}$$

Here $\chi_r(z,v) = \csc|\Theta_c(z,v)|$ if $|\Theta_c(z,v)| < \frac{\pi}{2}$ and $\csc|\Theta_c(z,v)| \leq e^r$. Otherwise set $\chi_r(z,v) = 1$.

Remark. Note that if $\Theta_{c_1}(z,v)$ gets smaller the upper bound in (53) increases. However once $|\Theta_{c_1}(z,v)|$ is small enough so that $\sin|\Theta_{c_1}(z,v)| \leq e^{-r}$ then the upper bound in (53) decreases sharply. In fact $\chi_r(z,v)$ takes values in the interval $[1,e^r]$.

Proof. Assume first that $\mathbf{t}_r(z,v) = 0$. Then by definition $z \in \mathcal{H}_c(1)$ for some $c \in \operatorname{Cusp}(G)$. Moreover we have $|\Theta_c(z,v)| < \frac{\pi}{2}$ and $\csc(|\Theta_c(z,v)|) \leq e^r$ (see (28)). In this case we have $\mathcal{K}_r(z,v) = e^{\mathbf{h}(z)}$ because z belongs to the 1-horoball at the cusp which absorbs $\gamma_{(z,v)}$. If $z \in \mathbf{Th}_G(10r)$ we have that $\mathcal{K}_r(z,v) \leq e^{10r}$ so the inequality (52) holds for such (z,v). If $(z,v) \in \mathcal{H}_c(10r)$ then (53) holds as well. From now on we assume that $\mathbf{t}_r(z,v) > 0$. Then $\mathcal{K}_r(z,v) \leq \iota(\gamma_{(z,v)}(\mathbf{t}_r(z,v)),\tau(G))$ by Proposition 7.1.

Let $(z,v) \in T^1\mathbf{H}$ and let $t_* = \min\{(3r+\mathbf{h}(z)),\mathbf{t}_r(z,v)\}$. By Proposition 5.6 we have that the segment $\gamma_{(z,v)}[t_*,\mathbf{t}_r(z,v)]$ is contained in $\mathbf{Th}_G(1)$ and therefore by Proposition 7.2 we have

$$\iota(\gamma_{(z,v)}[t_*,\mathbf{t}_r(z,v)],\tau(G)) \leq K(S)(\mathbf{t}_r(z,v)+1) \tag{54}$$

It remains to estimate $\iota(\gamma_{(z,v)}[0,t_*],\tau(G))$. Fix $0 \leq t < t_*$. Let $t_0 = 0$, $t_k = t$, and let $\{t_1,t_2,...,t_{k-1}\}$ be the ordered set of points where the segment $\gamma_{(z,v)}[0,t]$ intersects the 1-horocircles. Note that there can be at most $2t_0+2$ such segments, that is $k \leq 2t_0$. Each segment $\gamma_{(z,v)}[t_i,t_{i+1}]$ is either contained in $\mathbf{Th}_G(1)$ or in some $\mathcal{H}_c(1)$. If $\gamma_{(z,v)}[t_i,t_{i+1}]$ is contained in $\mathbf{Th}_G(1)$ then by Proposition 7.2 we have

$$\iota(\gamma_{(z,v)}[t_i,t_{i+1}],\tau(G)) \leq K(S)(|t_{i+1}-t_i|+1). \tag{55}$$

Assume that $\gamma_{(z,v)}[t_i, t_{i+1}]$ is contained in some $\mathcal{H}_c(1)$ (note that $\mathbf{a}_r(z,v) \neq c$). There are two cases. The first one is when $z \in \mathcal{H}_c(1)$ (then clearly $i = 0$). Then the entire segment $\gamma_{(z,v)}[t_0, t_1]$ is contained in $\mathcal{H}_c(1)$. Let $z_0 = \mathbf{z}_{\max}(\gamma_{(z,v)}, c)$. Then the segment $\gamma_{(z,v)}[t_0, t_1]$ can intersect at most $2N(S)e^{\mathbf{h}(z_0)}$ edges from $\lambda(G)$. On the other hand from (28) we have that $\mathbf{h}_c(z_0) - \mathbf{h}_c(z) \leq \log(\csc|\Theta_c(z,v)|)$ if $|\Theta_c(z,v)| \leq \frac{\pi}{2}$ and $\mathbf{h}_c(z_0) = \mathbf{h}_c(z)$ if $|\Theta_c(z,v)| > \frac{\pi}{2}$. This shows that

$$\iota(\gamma_{(z,v)}[t_0, t_1], \tau(G)) \leq 2N(S)\chi_r(z,v)e^{\mathbf{h}(z)}. \tag{56}$$

For any other segment $\gamma_{(z,v)}[t_i, t_{i+1}]$ that is contained in some $\mathcal{H}_c(1)$ we have that $\mathbf{h}_c(z_0) - 1 = \mathbf{h}_c(z_0) - \mathbf{h}_c(\gamma_{(z,v)}(t_i)) < r$ where $z_0 = \mathbf{z}_{\max}(\gamma_{(z,v)}, c)$. This follows from the fact that $\mathbf{a}_r(z,v) \neq c$, and from $\mathbf{h}_c(z_0) - \mathbf{h}_c(\gamma_{(z,v)}(t_i)) < \mathbf{h}_c(z_0) - \mathbf{h}_c(z)$. Then the segment $\gamma_{(z,v)}[t_i, t_{i+1}]$ can intersect at most $2N(S)e^{\mathbf{h}(z_0)}$ edges from $\lambda(G)$, that is

$$\iota(\gamma_{(z,v)}[t_i, t_{i+1}], \tau(G)) \leq 2N(S)e^{r+1}. \tag{57}$$

Since $t_0 \leq 3r + \mathbf{h}(z)$ and from (55), (56) and (57) we obtain

$$\iota(\gamma_{(z,v)}[0, t_*], \tau(G)) \leq 2N(S)\chi_r(z,v)e^{\mathbf{h}(z)} + (2(3r+\mathbf{h}(z))+1)2N(S)e^{r+1} + (2(3r+\mathbf{h}(z))+1)K(S),$$

(here we also used that there are at most $2t_0 + 2$ segments $\gamma_{(z,v)}[t_i, t_{i+1}]$).

Let $z \in \mathbf{Th}_G(10r)$. If $z \in \mathcal{H}_c(1)$ then from $1 \leq \chi_r(z,v) \leq r$ we obtain

$$\iota(\gamma_{(z,v)}[0, t_*], \tau(G)) \leq 2N(S)e^{11r} + (26r+1)2N(S)e^{r+1} + (26r+1)K(S).$$

Together with (54) this proves (52).

Suppose that $z \in \mathcal{H}_c(10r)$ for some $c \in \text{Cusp}(G)$. Then similarly from (55), (56) and (57) we get

$$\iota(\gamma_{(z,v)}[0, t_*], \tau(G)) \leq 2N(S)\chi_r(z,v)e^{\mathbf{h}(z)} + (2(3r+\mathbf{h}(z))+1)2N(S)e^{r+1} + (2(3r+\mathbf{h}(z))+1)K(S).$$

Since $\mathbf{h}(z) \geq 10r$ for r large enough we have

$$(2(3r + \mathbf{h}(z)) + 1)2N(S)e^{r+1} + (2(3r + \mathbf{h}(z)) + 1)K(S) \leq e^{\mathbf{h}(z)}.$$

Together with (54) this proves (53).

□

Proposition 7.4. *Let G be a normalised group and let $z \in \mathcal{H}_\infty(0)$, that is $y \geq 1$. Let $z_1, z_2 \in \mathbf{H}$, where $z_j = x_j + iy_j$, $j = 1, 2$, and let η be the geodesic segment between z_1 and z_2. Moreover let η_j be the vertical geodesic segment connecting z_j with the point $x_j + iy$ (that is η_j is orthogonal to the* 0*-horocircle at ∞). Then*

$$\iota(\eta, \tau(G)) \leq \iota(\eta_1, \tau(G)) + \iota(\eta_2, \tau(G)) + N(S)|x_1 - x_2|.$$

Proof. Let $\gamma \in \lambda(G)$. If γ intersects η then γ either intersects η_1 or η_2 or the horocyclic segment between the points $x_1 + iy$ and $x_2 + iy$. Since $y \geq 1$ and since G is normalised we have that if γ intersects this horocyclic segment then γ ends at ∞. On the other hand there are at most $N(S)|x_1 - x_2|$ such vertical geodesics in $\lambda(G)$. This proves the proposition.

□

7.1. **The proof of Lemma 5.2.** Let $t \geq 0$ and set

$$E_t = \{(z,v) \in T^1S : \mathbf{t}_r(z,v) \geq t\}.$$

We have

Proposition 7.5. *There exists $r_0 > 0$ so that for $r > r_0$ and every $t \geq r^2$ we have*

$$\Lambda(E_t) \leq e^{-q(S)\frac{t}{4}},$$

where $q(S) > 0$ is the constant from Lemma 5.1.

Proof. Fix $t \geq r^2$. We have

$$T^1S = T^1\mathbf{Th}_S(\frac{t}{2})\bigcup(T^1S \setminus T^1\mathbf{Th}_S(\frac{t}{2})).$$

One finds that

$$\Lambda\big((T^1S \setminus T^1\mathbf{Th}_S(\frac{t}{2}))\big) = \mathbf{n}e^{-\frac{t}{2}},$$

where $\mathbf{n}$ is the number of cusps in $\operatorname{Cusp}(S)$. Since $t \geq r^2$ for r large enough we have

$$\Lambda\big((T^1S \setminus T^1\mathbf{Th}_S(\frac{t}{2}))\big) \leq \frac{1}{2}e^{-\frac{t}{4}}. \tag{58}$$

It follows from Proposition 5.6 that for every $(z,v) \in T^1S$ and every $3r+\mathbf{h}(z) \leq t' \leq t$ we have that $\mathbf{g}_{t'}(z,v) \subset A(t-t')$, where $A(t)$ is the set defined in Lemma 5.1. If $(z,v) \in T^1\mathbf{Th}_S(\frac{t}{2})$ then

$$\mathbf{g}_{(3r+\frac{t}{2})}(z,v) \subset A(\frac{t}{2} - 3r).$$

It follows from Lemma 5.1 that

$$\Lambda\big(E_t \cap T^1\mathbf{Th}_S(\frac{t}{2})\big) = \Lambda\big(\mathbf{g}_{(3r+\frac{t}{2})}(E_t \cap T^1\mathbf{Th}_S(\frac{t}{2}))\big) \leq C(S)e^{-q(S)(\frac{t}{2}-3r)}.$$

Since $t \geq r^2$ for r large enough we have

$$\Lambda\big(E_t \cap T^1\mathbf{Th}_S(\frac{t}{2})\big) \leq \frac{1}{2}e^{-q(S)\frac{t}{4}}.$$

Together with (58) this proves the proposition. □

We are ready to prove Lemma 5.2. Let

$$T_r = \{(z,v) \in T^1S : \mathbf{h}(z) \geq 10r\}.$$

We have $T^1S \setminus \mathcal{D}_r(S) = E_{r^2} \cup T_r$. It follows from Proposition 7.3 and from the definition of $\varphi_r(z,v)$ that

$$\int\limits_{T^1S\setminus\mathcal{D}_r(S)} (\mathcal{K}_r(z,v) + \mathcal{K}_r(z,\omega v) + \mathcal{K}_r(z,\omega^2 v) + 1)\varphi_r(z)\, d\Lambda \leq \int\limits_{E_{r^2}\setminus T_r} D(S)e^{11r}\, d\Lambda +$$

$$+\int\limits_{T_r} D(S)e^{\mathbf{h}(z)}\chi_r(z,v)e^{(r-\frac{\mathbf{h}(z)}{2})}\, d\Lambda + \int\limits_{E_{r^2}} D(S)K(S)(\mathbf{t}_r(z,v)+1)\, d\Lambda +$$

$$+\int\limits_{T_r\setminus E_{r^2}} D(S)K(S)(\mathbf{t}_r(z,v)+1)\, d\Lambda.$$

We estimate each of the four integrals on the right hand side.

From Proposition 7.5 we have

$$\int_{E_{r^2}} D(S)e^{4r}\,d\Lambda = D(S)e^{4r}\Lambda(E_{r^2}) \leq D(S)e^{4r}e^{-q(S)\frac{r^2}{4}} \leq e^{-2r} = \mathbf{e}(2r),$$

for r large enough.

Next we estimate the second integral

$$\int_{T_r} D(S)e^{\mathbf{h}(z)}\chi(z,v)e^{(r-\frac{\mathbf{h}(z)}{2})}\,d\Lambda.$$

We have that $T_r = \cup_{c_i(S)\in \mathrm{Cusp}(S)} T^1\mathcal{H}_{c_i(S)}(10r)$. Let $c_i \in \mathrm{Cusp}(G)$ such that $[c_i]_G = c_i(S)$ and set $G = G_{c_i}$. Let $z \in \mathcal{H}_{c_i}(10r)$. Then $(z,v) = (x,y,\theta) = (z,\theta)$ in the polar coordinates on $T^1\mathbf{H}$. We have

$$\int_{-\pi}^{\pi} \chi_r(z,\theta)\,d\theta \leq Cr,$$

where C is some universal constant. This shows that for r large enough we have

$$\int_{T^1\mathcal{H}_{c_i}} D(S)e^{\mathbf{h}(z)}\chi(z,v)e^{(r-\frac{\mathbf{h}(z)}{2})}\,d\Lambda \leq Cr\int_{e^{10r}}^{\infty} \frac{D(S)ye^r}{\sqrt{y}}\,\frac{dy}{y^2} =$$

$$= D(S)Cre^r\int_{e^{10r}}^{\infty} y^{-\frac{3}{2}}\,dy \leq e^{-4r} < e^{-r}.$$

Next we estimate the third integral. Recall that if $f : X \to [0,\infty)$ is an integrable function on a measure space (X,μ) then

$$\int_X f\,d\mu = \int_0^{\infty} \mu(f^{-1}[t,\infty))\,dt.$$

Set $(X,\mu) = (E_{r^2},\Lambda)$ and $f(z,v) = \mathbf{t}_r(z,v)$. We find

$$\int_{E_{r^2}} \mathbf{t}_r(z,v)\,d\Lambda = \int_0^{\infty} \Lambda(E_t \cap E_{r^2})\,dt = r^2\Lambda(E_{r^2}) + \int_{r^2}^{\infty} \Lambda(E_t)\,dt.$$

This together with Proposition 7.5 gives

$$\int_{E_{r^2}} D(S)K(S)(\mathbf{t}_r(z,v)+1)\,d\Lambda \leq D(S)K(S)\Lambda(E_{r^2}) + D(S)K(S)\big(r^2\Lambda(E_{r^2}) + \int_{r^2}^{\infty} \Lambda(E_t)\,dt\big) \leq$$

$$\leq (1+r^2)D(S)K(S)e^{-q(S)\frac{r^2}{4}} + D(S)K(S)\int_{r^2}^{\infty} e^{-q(S)\frac{t}{4}}\,dt,$$

so for r large enough we have

$$\int_{E_{r^2}} D(S)K(S)(\mathbf{t}_r(z,v)+1)\,d\Lambda \leq e^{-2r} = \mathbf{e}(2r).$$

If $(z,v) \in T_r \setminus E_{r^2}$ then $(\mathbf{t}_r(z,v)+1) \leq r^2+1$. This implies that

$$\int\limits_{T_r \setminus E_{r^2}} D(S)K(S)(\mathbf{t}_r(z,v)+1)\,d\Lambda \leq D(S)K(S)(r^2+1)\Lambda(T_r) \leq r^3 e^{-10r} < e^{-r}.$$

We have estimated all four integrals and we find that

$$\int\limits_{T^1S \setminus \mathcal{D}_r(S)} (\mathcal{K}_r(z,v)+\mathcal{K}_r(z,\omega v)+\mathcal{K}_r(z,\omega^2 v)+1)\varphi_r(z)\,d\Lambda \leq P(r)e^{-r}.$$

This proves Lemma 5.2.

7.2. **The proof of Lemma 5.3.** From Proposition 5.16 we have

$$\int\limits_{H_r} (\mathcal{K}_r(z,v)+\mathcal{K}_r(z,\omega v)+\mathcal{K}_r(z,\omega^2 v)+1)\varphi_r(z)\,d\Lambda \leq$$

(59)

$$\leq \sum_{i=1}^{3} \int\limits_{\mathcal{E}_r^i(S)} (\mathcal{K}_r(z,v)+\mathcal{K}_r(z,\omega v)+\mathcal{K}_r(z,\omega^2 v)+1)\varphi_r(z)\,d\Lambda,$$

so in order to prove Lemma 5.3 we need to estimate from above the integrals on the right hand side.

We first estimate the integral

$$\int\limits_{\mathcal{E}_r^1(S)} (\mathcal{K}_r(z,v)+\mathcal{K}_r(z,\omega v)+\mathcal{K}_r(z,\omega^2 v)+1)\varphi_r(z)\,d\Lambda.$$

Let $(z,v) \in \mathcal{E}_r^1(S)$. Then $(z,v) \in \mathcal{D}_r(S)$. Let (z,v) also denote a lift of (z,v) to $\mathbf{H}$. Then $z \in \mathcal{H}_c(2r-\log 3)$, for some $c \in \operatorname{Cusp}(G)$ and c is an endpoint of $\mathbf{a}_r^1(z,v)$. Here $[c]_G = c_i(S)$ for some $c_i(S) \in \operatorname{Cusp}(S)$. Let $X_0(c_i(S))$ be the set of those $(z,v) \in \mathcal{E}_r^1(G)$ so that $z \in \mathcal{H}_c(2r-\log 3)$ and $\mathbf{a}_r(z,v)=c$. Let $X_1(c_i(S))$ be the set of those $(z,v) \in \mathcal{E}_r^1(G)$ so that $z \in \mathcal{H}_c(2r-\log 3)$ and $\mathbf{a}_r(z,\omega v)=c$. Then $\mathcal{E}_r^1(G) = X_0(c_i(S)) \cup X_1(c_i(S))$.

Let $(z,v) \in X_0(c_i(S))$. Since $(z,v) \in \mathcal{D}_r(S)$ we have $\max\{\mathbf{t}_r(z,v),\mathbf{t}_r(z,\omega v),\mathbf{t}_r(z,\omega^2 v)\} \leq r^2$. It follows from Proposition 5.2 that $|\Theta_c(z,\omega v)| \leq \pi e^{-r}$. This implies that $|\Theta_c(z,\omega^j v)| > \frac{\pi}{2}$, $j=1,2$. Since $\mathbf{a}_r(z,v)=c$ we have $\mathcal{K}_r(z,v)=e^{\mathbf{h}(z)}$. Combining this with (53) (recall the definition of $\chi_r(z,v)$) we get

$$\mathcal{K}_r(z,v)+\mathcal{K}_r(z,\omega v)+\mathcal{K}_r(z,\omega^2 v) \leq e^{\mathbf{h}(z)} + 2D(S)\big(e^{\mathbf{h}(z)}+K(S)(r^2+1)\big).$$

A single lift of the set $X_0(c_i(S))$ to $\mathbf{H}$ (with respect to the normalised group $G=G_c$) is contained in the set

$$\{(x,y,\theta) \in T^1\mathbf{H} :\, 0 \leq x \leq 1,\, \frac{e^{2r}}{3} \leq y \leq \infty,\, |\theta| \leq \pi e^{-r}\}.$$

Passing with the integration to the universal cover we get for r large that

$$\int\limits_{X_0(c_i(S))} (\mathcal{K}_r(z,v)+\mathcal{K}_r(z,\omega v)+\mathcal{K}_r(z,\omega^2 v)+1)\varphi_r(z)\,d\Lambda <$$

$$< \int_{-\pi e^{-r}}^{\pi e^{-r}} \left(\int_{\frac{e^{2r}}{3}}^{\infty} \big(y + 2D(S)\big(y + K(S)(1+r^2)\big) + 1\big) \frac{e^r}{\sqrt{y}} \frac{dy}{y^2} \right) d\theta \leq$$

$$\leq r^3 e^{-r} \int_{\frac{e^{2r}}{3}}^{\infty} e^r y^{-\frac{3}{2}} \, dy < 3r^3 e^{-r} = P(r) e^{-r}.$$

We repeat this for every $c_i(S) \in \operatorname{Cusp}(S)$. After repeating the same argument for $X_j(c_i(S))$, $j = 1, 2$, we obtain

$$\int_{\mathcal{E}_r^1(S)} (\mathcal{K}_r(z,v) + \mathcal{K}_r(z,\omega v) + \mathcal{K}_r(z,\omega^2 v) + 1)\varphi_r(z) \, d\Lambda < P(r)e^{-r}. \tag{60}$$

for r large enough.

It is somewhat more delicate to estimate the integrals over the sets $\mathcal{E}_r^2(S)$ and $\mathcal{E}_r^3(S)$. We need to prove several propositions first.

Proposition 7.6. *Let $s, t > 0$. Then for t large enough, we have*

$$I(s,t) = \int_{T^1\mathbf{Th}_S(s)} \iota(\gamma_{(z,v)}(t), \tau(S)) \, d\Lambda \leq t^2(s+t).$$

Proof. Denote the above integral by $I(s,t)$. Consider $\delta_{(z,v)}(t) = \iota(\gamma_{(z,v)}(t+1), \tau(S)) - \iota(\gamma_{(z,v)}(t), \tau(S))$. It follows from Proposition 7.2 that

$$\delta_{(z,v)}(t) \leq \max\{K(S), N(S)e^{\mathbf{h}(\gamma_{(z,v)}(t))}\}.$$

so we conclude that for any $(z,v) \in T^1 S$ we have

$$\delta_{(z,v)}(t) \leq K(S) + N(S)e^{\mathbf{h}(\gamma_{(z,v)}(t))}.$$

It follows that

$$I(t+1) - I(t) = \int_{T^1\mathbf{Th}_S(s)} \delta_{(z,v)}(t) \, d\Lambda \leq \int_{T^1\mathbf{Th}_S(s)} (K(S) + N(S)e^{\mathbf{h}(\gamma_{(z,v)}(t))}) \, d\Lambda.$$

Since $\mathbf{g}_t(T^1\mathbf{Th}_S(s)) \subset T^1\mathbf{Th}_S(s+t)$ and since $\mathbf{g}_t$ is measure preserving, we have

$$I(t+1) - I(t) \leq \int_{T^1\mathbf{Th}_S(s)} (K(S) + N(S)e^{\mathbf{h}(\gamma_{(z,v)}(t))}) \, d\Lambda = \int_{\mathbf{g}_t(T^1\mathbf{Th}_S(s))} (K(S) + N(S)e^{\mathbf{h}(z)}) \, d\Lambda \leq$$

$$\leq \int_{T^1\mathbf{Th}_S(s+t)} (K(S) + N(S)e^{\mathbf{h}(z)}) \, d\Lambda = \int_{T^1\mathbf{Th}_S(0)} (K(S) + N(S)e^{\mathbf{h}(z)}) \, d\Lambda +$$

$$+ \sum_{i=1}^{n} \int_{H(c_i(S))} (K(S) + N(S)e^{\mathbf{h}(z)}) \, d\Lambda,$$

where $H(c_i(S)) = T^1\mathcal{H}_{c_i(S)}(0) \setminus T^1\mathcal{H}_{c_i(S)}(s+t)$. We have

$$\int_{T^1\mathbf{Th}_S(0)} (K(S) + N(S)e^{\mathbf{h}(z}) \, d\Lambda \leq \int_{T^1\mathbf{Th}_S(0)} (K(S) + N(S)) \, d\Lambda \leq C,$$

where $C > 0$ depends only on S. On the other hand, by passing to the group $G = G_{c_i(S)}$ and using the expression for the volume element $d\Lambda$ we have

$$\int\limits_{H(c_i(S))} (K(S)+N(S)e^{\mathbf{h}(z)})\, d\Lambda = \int\limits_{T^1\mathcal{H}_\infty(0)\setminus T^1\mathcal{H}_\infty(s+t)} (K(S)+N(S)e^{\mathbf{h}_\infty(z)})\, d\Lambda =$$

$$= \int\limits_0^1 \left(\int\limits_1^{e^{s+t}} \frac{K(S)+N(S)y}{y^2}\, dy \right) dx \leq K(S) + N(S)(s+t).$$

Combining these estimates we have $I(t+1) - I(t) \leq C + K(S) + N(S)(s+t)$ and therefore $I(t) \leq t(C + K(S) + N(S)(s+t))$. For t large enough we have $I(t) \leq t^2(t+s)$ which proves the lemma. □

We have already discussed (in Section 6) the cross sectional area for two dimensional subsets of $T^1\mathbf{H}$ with respect to the equidistant flow $\mathcal{S}_t$.

Definition 7.2. *Let $E \subset \mathbf{H}$ be a domain, and let $u(z)$ be a smooth unit vector field on E. Then $E(u) = \{(z, u(z)) : z \in E\}$ is a surface sitting inside the three dimensional manifold $T^1\mathbf{H}$. Let $0 \leq t_1 \leq t_2 \leq \infty$. Define $\mathcal{U}(E(u), t_1, t_2) = \cup_{t_1 \leq s \leq t_2} \mathbf{g}_s(E(u))$. For simplicity, set $\mathcal{U}(E(u), 0, t) = \mathcal{U}(E(u), t)$. We use the same notation for $E \subset S$.*

We say that a unit vector field u defined on a domain $E \subset \mathbf{H}$ is transverse to E if $\mathbf{g}_s(E(u)) \cap E(u) = \emptyset$ for every $s > 0$. Assume that u is transverse to E. Since $\mathbf{g}_t$ preserves the volume, we have that the quotient $\Lambda(\mathcal{U}(E(u), t))/t$ does not depend on $t \geq 0$. We call $\Lambda(\mathcal{U}(E(u), t))/t$ the cross sectional area of E with respect to $u(z)$. In fact, the flow $\mathbf{g}_t$ induces an area form $d\eta(u)$ on E so that the cross sectional area of E agree with the area of E with respect to $d\eta(u)$. The two form $d\eta(u)$ is obtained by contracting $|d\Lambda|$ by the vector field $u(z)$. One can verify that since each vector $u(z)$ has the unit length, and since $d\Lambda = y^{-2}\, dx \wedge dy \wedge d\theta$ we have that $d\eta(u) = \sigma(z)\, dx \wedge dy$ where $0 \leq \sigma(z) \leq y^{-2}$. That is the density of the two form $d\eta(u)$ is always bounded above by the density of the hyperbolic metric.

Let $\psi = \psi_{E(u)} : \mathcal{U}(E(u), \infty) \to E(u)$ be the projection map, that is on each slice $\mathbf{g}_s(E(u))$, $0 \leq s < \infty$, the map ψ agrees with $(\mathbf{g}_s)^{-1}$. Since u is transverse to E the map ψ is well defined.

Let $f_t : E(u) \to \mathbf{R}$, $t_1 \leq t \leq t_2$, and $f : \mathcal{U}(E(u), t) \to \mathbf{R}$ be integrable functions. If for every such t we have $f(z, v) \leq f_t(\psi_{E(u)}(z, v))$ for every $(z, v) \in \mathbf{g}_t(E(u))$ then

$$\int\limits_{\mathcal{U}(E(u), t_1, t_2)} f\, d\Lambda \leq \int\limits_{t_1}^{t_2} \left(\int\limits_{E(u)} f_t d\eta(u) \right) dt, \tag{61}$$

and if $f_t(\psi_{E(u)}(z, v)) \leq f(z, v)$ for every $(z, v) \in \mathbf{g}_t(E(u))$ then

$$\int\limits_{t_1}^{t_2} \left(\int\limits_{E(u)} f d\eta(u) \right) dt \leq \int\limits_{\mathcal{U}(E(u), t)} f\, d\Lambda. \tag{62}$$

We are particularly interested in the vector fields $u_1(z)$ and $u_2(z)$. The vector $u_1(z)$ is the unique vector such that the point $(z, u_1(z)) \in T^1\mathbf{H}$ corresponds to the coordinates $(x, y, -\frac{\pi}{2})$. Observe that u_1 is transverse to $\mathbf{H}$. The vector $u_2(z)$ is the unique vector such that the point $(z, u_2(z)) \in T^1\mathbf{H}$ corresponds to the coordinates $(x, y, -\pi)$. One can verify that $d\eta(u_1) = y^{-2}\,dxdy$ and that $d\eta(u_2) = 0$. It is not surprising that $d\eta(u_2) = 0$ since the corresponding set $\mathcal{U}(E(u_2), t)$ is two dimensional for any set $E \subset \mathbf{H}$ and therefore the cross sectional area of E with respect to $u_2(z)$ is equal to zero.

Fix $c_i(S) \in \operatorname{Cusp}(S)$. For $y \geq 1$ set

$$A_j(t, y, c_i(S)) = \int_0^1 \iota(\gamma_{(z,u_j(z))}(t), \tau(G_{c_i(S)}))\,dx,$$

where $j = 1, 2$. Set $G = G_{c_i(S)}$. Note that if $z \in \gamma \cap \mathcal{H}_\infty(0)$ and $\gamma \in \lambda(G)$ then γ has ∞ as its endpoint. This implies that for $1 \leq y$ we have

$$A_2(t, 1, c_i(S)) = A_2(t + \log y, y, c_i(S)), \tag{63}$$

and $A_2(t, y, c_i(S)) = 0$ for $t \leq \log y$.

Fix $y \geq 1$. For every $t \geq 0$ the set of points $\gamma_{(z,u_1(z))}(t)$ is a horocircle in $\mathbf{H}$ that bounds a horoball at ∞ (recall $z = x + iy$). Moreover, there exist numbers $d(t) \geq 0$ and $r(t) \geq 0$ (that depend only on t), such that for $z' = z + yd(t)$ we have $\gamma_{(z,u_1(z))}(t) = \gamma_{(z',u_2(z'))}(r(t))$. The functions $d(t)$ and $r(t)$ are increasing in t. It is elementary to verify that $0 \leq d(t) \leq 1$ and $0 \leq t - r(t) < 1$.

We apply Proposition 7.4 to the segment $\gamma_{(z,u_1(z))}[0, t]$. We have

$$\iota(\gamma_{(z,u_1(z))}(t), \tau(G)) \leq \iota(\gamma_{(z',u_2(z'))}(r(t)), \tau(G)) + N(S)yd(t).$$

This shows that for every $t \geq 0$ we have

$$\begin{aligned}\iota(\gamma_{(z',u_2(z'))}(t-1), \tau(G)) &\leq \iota(\gamma_{(z',u_2(z'))}(r(t)), \tau(G)) \leq \iota(\gamma_{(z,u_1(z))}(t), \tau(G)) \leq \\ &\iota(\gamma_{(z',u_2(z'))}(r(t)), \tau(G)) + N(S)y \leq \iota(\gamma_{(z',u_2(z'))}(t), \tau(G)) + N(S)y,\end{aligned}$$

that is

$$\iota(\gamma_{(z',u_2(z'))}(t-1), \tau(G)) \leq \iota(\gamma_{(z,u_1(z))}(t), \tau(G)) \leq \iota(\gamma_{(z',u_2(z'))}(t), \tau(G)) + N(S)y.$$

Since the vector $u_2(z')$ is obtained by translating the vector $u_2(z)$ for $yd(t)$ and $d(t)$ does not depend on x we get

$$A_2(t-1, y, c_i(S)) \leq A_1(t, y, c_i(S)) \leq A_2(t, y, c_i(S)) + N(S)y. \tag{64}$$

Proposition 7.7. *For $y \geq 1$ and any $c_i(S)$ we have*

$$A_2(t, y, c_i(S)) \leq (t + 3)^4,$$

and

$$A_1(t, y, c_i(S)) \leq (t + 3)^4 + N(S)y,$$

for t large enough.

Proof. Fix $c_i(S) \in \operatorname{Cusp}(S)$ and let $G = G_{c_i(S)}$. Let $E \subset \mathbf{H}$ be the set given by $E = \{z : 0 \leq x \leq 1, \text{ and } 4 \leq y \leq 5\}$. Let $\mathcal{U}(E(u_1), 1) \subset T^1\mathbf{H}$ be the corresponding set (see the above definition). Note that $\mathcal{U}(E(u_1), 1) \subset T^1\mathbf{Th}_G(\log 5)$. Moreover, the set $\mathcal{U}(E(u_1),1)$ injects into $T^1\mathbf{Th}_S(\log 5)$ under the standard covering map. Therefore, from Proposition 7.6 for t large enough we have

$$\int\limits_{\mathcal{U}(E(u_1),1)} \iota(\gamma_{(z,v)}(t),\tau(G))\,d\Lambda \leq \int\limits_{T^1\mathbf{Th}_S(\log 5)} \iota(\gamma_{(z,v)}(t),\tau(S))\,d\Lambda =$$

(65)

$$= \int\limits_{T^1\mathbf{Th}_S(\log 5)} \iota(\gamma_{(z,v)}(t),\tau(S))\,d\Lambda \leq t^2(\log 5+t) < t^2(2+t).$$

Let $\psi:\mathcal{U}(E(u_1),1)\to E(u_1)$ be the restriction of the projection map $\psi_{E(u_1)}:\mathcal{U}(E(u_1),\infty)\to E(u_1)$ introduced above. It holds that $\iota(\gamma_{\psi(z,v)}(t),\tau(G))\leq \iota(\gamma_{(z,v)}(t),\tau(G))+C_1$ for every $t>0$ where C_1 is a constant that depends only on S. The constant C_1 bounds above the number of intersections between the geodesic segment $\gamma_{(w,u_1(w))}[0,1]$ and $\tau(G)$ when $w\in E$. From (62) we have

$$\int\limits_E \iota(\gamma_{(z,u_1(z))}(t),\tau(G))d\eta(u_1) \leq \int\limits_{\mathcal{U}(E(u_1),1)} (\iota(\gamma_{(z,v)}(t),\tau(G))+C_1)\,d\Lambda.$$

Combining this with (65) yields

$$\int\limits_E \iota(\gamma_{(z,u_1(z))}(t),\tau(G))d\eta(u_1) \leq (t+1)^2(3+t)+C_1\Lambda(\mathcal{U}(E(u_1),1)) \leq 2t^3,$$

for t large enough. Since $d\eta(u_1)=y^{-2}\,dxdy$ this gives

$$\int\limits_4^5 A_1(t,y,c_i(S))\,dy \leq 25\int\limits_4^5 A_1(t,y,c_i(S))y^{-2}\,dy =$$

(66)

$$=25\int\limits_4^5\int\limits_E \iota(\gamma_{(z,u_1(z))}(t),\tau(G))d\eta(u_1)\leq t^4.$$

From (64) and (66) we obtain

$$\int\limits_4^5 A_2(t,y,c_i(S))\,dy \leq (t+1)^4.$$

Observe that for $y\in[4,5]$ we have $A_2(t,1,c_i(S))\leq A_2(t+\log y,y,c_i(S))\leq A_2(t+2,y,c_i(S))$. Therefore

$$A_2(t,1,c_i(S)) \leq \int\limits_4^5 A_2(t+2,y,c_i(S))\,dy \leq (t+3)^4.$$

Combining this with (63) we conclude that for every $y\geq 1$ we have $A_2(t,y,c_i(S))\leq A_2(t,1,c_i(S))\leq (t+3)^4$. The second estimate in the proposition follows from the first one, and from (64).

□

Remark. If instead of $u_1(z)$ we use $-u_1(z)$ to define $A_1(t,y,c_i(S))$ that is

$$\widehat{A}_1(t,y,c_i(S)) = \int\limits_0^1 \iota(\gamma_{(z,-u_1(z))}(t), \tau(G_{c_i(S)}))\,dx.$$

we obtain the same estimate $\widehat{A}_1(t,y,c_i(S)) \leq (t+3)^4 + N(S)y$.

Again fix $c_i(S) \in \operatorname{Cusp}(S)$. Let $z_t = \gamma_{(z,u_1(z))}(t)$ and $v_t = \gamma'_{(z,u_1(z))}(t)$. For $y \geq 1$ set

$$I(t,y,c_i(S)) = \int\limits_0^1 \left(\iota(\gamma_{(z,u_1(z))}(t), \tau(G_{c_i})) + \sum_{j=1}^{2} \iota(\gamma_{(z_t,\omega^j(-v_t))}(r^2+1), \tau(G_{c_i})) \right) dx.$$

Set $G = G_{c_i}(S)$. We want to estimate $I(t,y,c_i(S))$. For every $t \geq 0$ and $y \geq 1$ the set of points $\gamma_{(z_t,\omega^j(-v_t))}(r^2+1)$ is a horocircle in $\mathbf{H}$ that bounds a horoball at ∞ (recall $z = x+iy$). Since $0 < |\Theta_\infty(z_t,(-v_t))| \leq \frac{\pi}{2}$ we see that $|\Theta_\infty(z_t,\omega^j(-v_t))| \geq \frac{\pi}{6}$ for any value of t. This implies that the point $\gamma_{(z_t,\omega^j(-v_t))}(r^2+1)$ lies below the the horocircle $\partial\mathcal{H}_\infty(\log y)$ for r large enough.

Fix $y \geq 1$. Recall the functions $d(t)$ and $r(t)$ defined above (and the corresponding point $z' = z+yd(t)$). Similarly, we define the functions $d_j(t)$ and $r_j(t)$ as follows. Let $z''_j \in \mathbf{H}$ so that z''_j lies on the same horocircle $\partial\mathcal{H}_\infty(\log y)$ as the points z and z', and so that $\gamma_{(z''_j,u_2(z''_j))}(r_j(t)) = \gamma_{(z_t,\omega^j(-v_t))}(r^2+1)$. Then $z''_j = z + yd_j(t)$. The functions $d_j(t)$ and $r_j(t)$ are increasing in t and do not depend on z. It is elementary to verify that $-1 < d_j(t) < 2$ and $0 \leq r(t) < t + r^2 + 1$.

As we explained (and stated) above from Proposition 7.4 we have

$$\iota(\gamma_{(z,u_1(z))}(t), \tau(G)) \leq \iota(\gamma_{(z',u_2(z'))}(r(t)), \tau(G)) + N(S)y.$$

Similarly we apply Proposition 7.4 to the segment $\gamma_{(z_t,\omega^j(-v_t))}(r^2+1)$ and get

$$\iota(\gamma_{(z_t,\omega^j(-v_t))}(r^2+1), \tau(G)) \leq \iota(\gamma_{(z',u_2(z'))}(r(t)), \tau(G)) + \iota(\gamma_{(z''_j,u_2(z''_j))}(r_j(t)), \tau(G)) + 2N(S)y,$$

(here we used that the Euclidean length of the horocycilic segment between the points z' and z''_j is at most 2). This yields that

$$I(t,y,c_i(S)) \leq 3A_2(t,y,c_i(S)) + 2A_2(t+r^2+1,y,c_i(S)) + 5N(S)y,$$

so from Proposition 7.7, for r large enough we get

$$I(t,y,c_i(S)) \leq 3(t+3)^4 + 2(t+r^2+3)^4 + 5N(S)y,$$

that is

$$I(t,y,c_i(S)) < 3(t+r^2+3)^4 + 5N(S)y. \tag{67}$$

Remark. If instead of $u_1(z)$ we use $-u_1(z)$ to define $I(t,y,c_i(S))$, that is

$$I(t,y,c_i(S)) = \int\limits_0^1 \left(\iota(\gamma_{(z,-u_1(z))}(t), \tau(G_{c_i})) + \sum_{j=1}^{2} \iota(\gamma_{(z_t,\omega^j(-v_t))}(r^2+1), \tau(G_{c_i}) \right) dx.$$

we get that the same estimate (67) holds.

We are ready to estimate the integral

$$\sum_{i=2}^{3} \int\limits_{\mathcal{E}_r^i(S)} (\mathcal{K}_r(z,v)+\mathcal{K}_r(z,\omega v)+\mathcal{K}_r(z,\omega^2 v)+1)\varphi_r(z)\, d\Lambda.$$

Note that for $c_i(S) \in \operatorname{Cusp}(S)$ the vector fields $u_j(z)$ are well defined on $T^1\mathcal{H}_{c_i(S)}(0) \subset T^1 S$.

Let $E = \mathbf{Th}_S(12r) \setminus \mathbf{Th}_S(0)$. Let $(z,v) \in \mathcal{E}_r^2(G)$. We also use (z,v) to denote a lift of (z,v) to $\mathbf{H}$. Then

$$r - r^2 e^{-r} < \mathbf{h}_{\max}(\mathbf{z}_{\max}(\gamma_{(z,v)}, c)) - \mathbf{h}_{\max}(z) < r + r^2 e^{-r},$$

and $\mathbf{z}_{\max}(\gamma_{(z,v)}, c) \in \mathcal{H}_\infty(1-r^2e^{-r})$ for some cusp $c \in \operatorname{Cusp}(G)$. Set $G = G_c$. It follows from (31) that for r large enough, we have $\mathbf{z}_{\max}(\gamma_{(z,v)}, c)) = \gamma_{(z,v)}(t)$ for some $(r - r^2e^{-r}) + \log 2 - e^{-r} < t < (r + r^2 e^{-r}) + \log 2 + e^{-r}$. Combining this with the fact that $z \in \mathbf{Th}_G(10r)$ (recall that $E_r^2(S) \subset \mathcal{D}_r(S) \subset T^1\mathbf{Th}_S(10r)$), this implies that $(z,-v) \in \mathcal{U}(E(u_1^*), r+\log 2 - 2r^2e^{-r}, r + \log 2 + 2r^2 e^{-r})$ where u_1^* is either equal to u_1 or to $-u_1$.

Recall that if $(z,v) \in \mathcal{E}_r^2(S)$ then (z,v) does not belong to $\mathcal{E}_r^1(S)$ so for $j=0,1,2$, by Proposition 7.1 we have

$$\mathcal{K}_r(z,\omega^j v) \le \iota(\gamma_{(z,\omega^j v)}(r^2+1), \tau(S)).$$

In particular for $(z,v) \in \mathbf{g}_t(E(u_1^*))$ we have the estimate

$$\mathcal{K}_r(z,v) \le \iota(\gamma_{(z,-v)}(r^2+1),\tau(S)) < \iota(\gamma_{(z,(-v))}(t),\tau(S)) + \iota(\gamma_{-\psi(z,v)}(r^2+1),\tau(S)),$$

where $\psi : \mathcal{U}(E(u_1^*), (r+\log 2 - 2r^2e^{-r}), (r+\log 2 + 2r^2e^{-r})) \to E(u_1^*)$ is the projection map. Define $f : \mathcal{U}(E(u_1^*), (r+\log 2 - 2r^2e^{-r}), (r+\log 2+2r^2e^{-r})) \to \mathbf{R}$ by

$$f(z,v) = 1 + \sum_{i=0}^{2} \mathcal{K}_r(z,\omega^j v),$$

and $f_t : E(u_1^*) \to \mathbf{R}$ by

$$f_t(z) = 1 + (\iota(\gamma_{(z,u_1(z))}(t),\tau(S)) + \iota(\gamma_{(z,-u_1(z))}(r^2+1),\tau(S)) +$$
$$+\iota(\gamma_{(z_t,\omega(-v_t))}(r^2+1),\tau(S)) + \iota(\gamma_{(z_t,\omega^2(-v_t))}(r^2+1),\tau(S)).$$

We apply (61) separately for $u_1^* = u_1$ and $u_1^* = -u_1$. Adding the two inequalities, we obtain

$$\int\limits_{\mathcal{E}_r^2(S)} (\mathcal{K}_r(z,v)+\mathcal{K}_r(z,\omega v)+\mathcal{K}_r(z,\omega^2 v)+1)\varphi_r(z)\, d\Lambda < \int\limits_{\mathcal{E}_r^2(S)} (\mathcal{K}_r(z,v)+\mathcal{K}_r(z,\omega v)+\mathcal{K}_r(z,\omega^2 v)+1)\, d\Lambda \le$$

$$2 \int\limits_{r+\log 2 - 2r^2e^{-r}}^{r+\log 2+2r^2e^{-r}} \left(\int\limits_E f_t d\eta(u_1) \right) dt.$$

Passing to the universal cover for each cusp $c_i(S)$ and applying the Fubini theorem to the last integral, yields that (recall $\widehat{A}_1(t,y,c_i(S))$ from the remark after the proof of Proposition 7.7)

$$\int_{E} f_t d\eta(u_1) \leq \mathbf{n} \int_{1}^{e^{12r}} \Big(1 + I(t,y,c_i(S)) + \widehat{A}_1(r^2+1,y,c_i(S))\Big)\, y^{-2}\, dy <$$

$$< \mathbf{n} \int_{1}^{e^{12r}} \big(1 + (t+r^2+3)^4 + 5N(S) + (t+3)^4 + N(S)\big)\, \frac{dy}{y},$$

where $\mathbf{n}$ is the number of cusps in $\text{Cusp}(S)$. By using (67) and Proposition 7.7 and since $t \leq r+1$ we have

$$\int_{\mathcal{E}^2_r(S)} (\mathcal{K}_r(z,v) + \mathcal{K}_r(z,\omega v) + \mathcal{K}_r(z,\omega^2 v) + 1)\varphi_r(z)\, d\Lambda <$$

(68)

$$< \int_{r+\log 2 - 2r^2e^{-r}}^{r+\log 2+2r^2e^{-r}} \left(\int_{1}^{12r} \big(1 + (t+r^2+3)^4 + 5N(S) + (t+3)^4 + N(S)\big)\frac{dy}{y} \right) dt \leq P(r)e^{-r}.$$

for r large enough.

Let $E = \mathbf{Th}_S(1+r^2e^{-r}) \setminus \mathbf{Th}_S(1-r^2e^{-r})$. Recall that if $(z,v) \in \mathcal{E}^3_r(G)$ then for some $r-1 < t < r^2+1$ and some $c \in \text{Cusp}(G)$ we have that $\mathbf{z}_{\max}(\gamma_{(z,v)}, c) = \gamma_{(z,v)}(t)$ and

$$1 - r^2e^{-r} < \mathbf{h}_c(\mathbf{z}_{\max}(\gamma_{(z,v)}, c)) < 1 + r^2e^{-r}.$$

This implies that if $(z,v) \in \mathcal{E}^3_r(S)$ then $(z,-v) \in \mathcal{U}(E(u_1^*),\, (r-1),\, (r^2+1))$ where u_1^* is either equal to u_1 or to $-u_1$.

Recall that if $(z,v) \in \mathcal{E}^3_r(S)$ then (z,v) does not belong to $\mathcal{E}^1_r(S)$, for $j = 0,1,2$. By Proposition 7.1 we have

$$\mathcal{K}_r(z,\omega^j v) \leq \iota(\gamma_{(z,\omega^j v)}(r^2+1), \tau(S)).$$

In particular, for $(z,-v) \in \mathbf{g}_t(E(u_1^*))$ we have the estimate

$$\mathcal{K}_r(z,v) \leq \iota(\gamma_{(z,v)}(r^2+1), \tau(S)) < \iota(\gamma_{(z,v)}(t), \tau(S)) + \iota(\gamma_{-\psi(z,v)}(r^2+1), \tau(S)),$$

where $\psi : \mathcal{U}(E(u_1^*),\, (r-1),\, (r^2+1)) \to E(u_1^*)$ is the projection map. Define the maps $f : \mathcal{U}(E(u_1^*),\, (r-1),\, (r^2+1)) \to \mathbf{R}$ and $f_t : E(u_1^*) \to \mathbf{R}$ as above. We apply (62) separately for $u_1^* = u_1$ and $u_1^* = -u_1$. Adding the two inequalities we obtain

$$\int_{\mathcal{E}^3_r(S)} (\mathcal{K}_r(z,v)+\mathcal{K}_r(z,\omega v)+\mathcal{K}_r(z,\omega^2 v)+1)\varphi_r(z)\, d\Lambda < \int_{\mathcal{E}^3_r(S)} (\mathcal{K}_r(z,v)+\mathcal{K}_r(z,\omega v)+\mathcal{K}_r(z,\omega^2 v)+1)\, d\Lambda \leq$$

$$2 \int_{r-1}^{r^2+1} \left(\int_{E} f_t d\eta(u_1) \right) dt.$$

Passing to the universal cover for each cusp $c_i(S)$ and applying the Fubini theorem to the last integral, yields that

$$\int_E f_t d\eta(u_1) \leq \mathbf{n} \int_{e^{1-r^2e^{-r}}}^{e^{1+r^2e^{-r}}} \big(1 + I(t,y,c_i(S)) + \widehat{A}_1(r^2+1,y,c_i(S))\big) y^{-2}\, dy,$$

where $\mathbf{n}$ is the number of cusps in $\operatorname{Cusp}(S)$. By using (67) and Proposition 7.7 we have

$$\int_{\mathcal{E}_r^3(S)} (\mathcal{K}_r(z,v) + \mathcal{K}_r(z,\omega v) + \mathcal{K}_r(z,\omega^2 v) + 1)\varphi_r(z)\, d\Lambda <$$

(69)

$$< 2\mathbf{n} \int_{r-1}^{r^2+1} \left(\int_{e^{1-r^2e^{-r}}}^{e^{1+r^2e^{-r}}} \big(1 + (t+r^2+3)^4 + 5N(S) + (t+3)^4 + N(S)\big) \frac{dy}{y} \right) dt = \mathbf{e}(r),$$

since $t \leq r^2+1$.

Putting together (60), (68), and (69), and replacing these in (59), we get

$$\int_{H_r} (\mathcal{K}_r(z,v) + \mathcal{K}_r(z,\omega v) + 1)\varphi_r(z)\, d\Lambda = \mathbf{e}(r).$$

which proves Lemma 5.3.

8. Appendix

To prove Lemma 5.1 it is enough to prove the following somewhat more general theorem.

Theorem 8.1. *Let S be a hyperbolic non-compact finite-area Riemann surface and let $K \subset S$ be compact. Then we can find $C, q > 0$ such that*

$$A_t = \{(z,v) \in T^1S : \gamma_(z,v)[0,t] \subset K\}.$$

satisfies $\Lambda(A_t(K)) \leq Ce^{-qt}$ for any $t \geq 0$.

Proof. We may assume that $K = \mathbf{Th}_S(h_0)$ for some $h_0 > 0$. We fix a proper ideal triangulation τ for S, and let $\lambda = \lambda(\tau)$ denote the set of geodesic edges of τ. For any $(z,v) \in T^1S$, we can record the sequence of left and right turns taken by the geodesic ray $\gamma_{(z,v)}[0,\infty)$ to obtain a sequence $R^{a_1}L^{a_2}\ldots$, where $a_i = a_i(z,v) \in \mathbf{N}$. This sequence is finite if and only if $\mathbf{p}(z,v) \in \operatorname{Cusp}(G)$. We let $(s_i)_{i=0}^{\infty}$ be the times such that $\gamma_{(z,v)}(s_i) \subset \lambda$, and we observe that we can find $C_1 > C_0 > 0$ ($C_i = C_i(S, \mathbf{Th}_S(h_0))$) such that $C_0 < s_{i+1} - s_i < C_1$ as long as $\gamma_{(z,v)}[s_i, s_{i+1}] \subset \mathbf{Th}_S(h_0)$. Moreover, if $\gamma_{(z,v)}[0, s(\sum_{i=1}^k a_i)] \subset \mathbf{Th}_S(h_0)$ (where $s(i) = s_i$), then $a_i \leq B$ for $i = 1, \ldots, k$, where $B \in \mathbf{N}$, depends only on $\mathbf{Th}_S(h_0)$.

We let

$V_n(B) = \{(z,v) \in \mathbf{Th}_S(h_0) : a_1, \ldots, a_{n+1}$ is well-defined, and $a_i \leq B$ for $i = 1, \ldots n\}$.

We will show that there are $C, q > 0$ such that $\Lambda(V_n(B)) \leq Ce^{-qn}$; this will complete the proof of the theorem.

To this end, we fix a triangle $T \in \tau_G$ (τ_G is the lift of τ to $\mathbf{H}$). We observe that if $(z,v) \in T$, then the sequence $(a_i(z,v))$ depends only on $\mathbf{p}(z,v)$. We let $W_n(B) = \mathbf{p}(V_n(B))$. For any $(b_i)_{i=1}^k$, where b_i are positive integers, the set

$$\{\mathbf{p}(z,v) : (a_i(z,v))_{i=1}^{k+1} \text{ is well-defined, and } a_i = b_i \text{ for } 1 \le i \le k\}$$

is an open interval in $\partial\mathbf{H}$ whose endpoints are joined by an element of λ_G. Therefore $W_n(B)$ is a disjoint union of B^n such intervals. For any such interval $I \subset W_n(B)$, we let $J \subset U$ be the least interval such that $I \cap W_{n+2}(B) \subset J$. The following two facts are central to our argument:

(1) The closure of J is a compact subset of I.
(2) The cross ratio $R(J,I)$ can take on only finitely many values (depending only on τ and B).

Here

$$R(J,I) \equiv \frac{(j_1 - j_0)(i_1 - i_0)}{(j_0 - i_0)(i_1 - j_1)},$$

where $I = (i_0, i_1)$ and $J = (j_0, j_1)$, is a Möbius invariant of (J,I). It follows that

$$\frac{|J|}{|I|} \le \eta(\tau, B) < 1,$$

where we fix a unit disk model $\mathbf{D}$ for $\mathbf{H}$, and let $|J|$ be the arc length for $J \subset \partial\mathbf{D}$. Therefore

$$|W_{n+2}(B) \cap I| \le \eta(\tau,B)|I|,$$

so

$$|W_{2n}(B)| \le 2\pi(\eta(\tau,B))^n = 2\pi e^{-nq},$$

where $q = -\log \eta(\tau,B) > 0$. Since for any interval $A \subset \partial\mathbf{D}$ we have

$$\Lambda(\{(z,v) \in \mathbf{Th}_G(h_0) \cap T : \mathbf{p}(z,v) \in A\}) \le C(\mathbf{Th}_S(h_0), \tau)|A|,$$

we find that

$$|V_{2n}(B)| \le Ce^{-nq},$$

where C depends only on $\mathbf{Th}_S(h_0)$ and τ.

□

References

[1] L. Ahlfors, *Lectures on quasiconformal mappings.* Second edition. With supplemental chapters by C. J. Earle, I. Kra, M. Shishikura and J. H. Hubbard. University Lecture Series, 38. American Mathematical Society, Providence, RI, (2006).

[2] I. Biswas, S. Nag, D. Sullivan, *Determinant bundles, Quillen metrics and Mumford isomorphisms over the universal commensurability Teichmller space.* Acta Math. 176 no. 2, 145-169 (1996)

[3] L. Chekhov, R. Penner, *On quantizing Teichmüller and Thurston theories.* Handbook of Teichmller theory. Vol. I, 579–645, IRMA Lect. Math. Theor. Phys., 11, Eur. Math. Soc., Zrich, (2007)

[4] L. Ehrenpreis, *Cohomology with bounds.* Symposia Mathematics IV, Academic Press, 389–395, (1970)

[5] M. Linch, *A comparison of metrics on Teichmller space.* Proc. Amer. Math. Soc. 43, 349–352 (1974)

[6] V. Markovic, D. Šarić, *The Teichmüller distance between finite index subgroups of* $\mathbf{PSL}(2,\mathbf{Z})$. Preprint.

[7] V. Markovic, D. Šarić, *Teichmller mapping class group of the universal hyperbolic solenoid.* Trans. Amer. Math. Soc. 358 no. 6 2637–2650 (2006)

[8] J. Parkkonen, V. Ruuska, *Finite degree holomorphic covers of compact Riemann surfaces.* Acta Math. Sinica 23 no. 1, 89–94. (2007)

[9] R. Penner, D. Saric, *Teichmuller theory of the punctured solenoid.* arXiv:math/0508476

Mathematics Department
Stony Brook University
Stony Brook, 11794
NY, USA
and
University of Warwick
Institute of Mathematics
Coventry, CV4 7AL, UK

E-mail address: `kahn@math.sunysb.edu, v.markovic@warwick.ac.uk`