

**RESEARCH ARTICLE**

# Counting curves on orbifolds

Viveka Erlandsson<sup>1,2</sup> | Juan Souto<sup>3</sup>

<sup>1</sup>School of Mathematics, University of Bristol, Bristol, UK

<sup>2</sup>Department of Mathematics and Statistics, UiT The Arctic University of Norway, Tromsø, Norway

<sup>3</sup>Université de Rennes, CNRS, Rennes, France

**Correspondence**

Viveka Erlandsson, School of Mathematics, University of Bristol, Bristol BS8 1UG, UK.  
Email: [v.erlandsson@bristol.ac.uk](mailto:v.erlandsson@bristol.ac.uk)

**Funding information**

Academy of Finland, Grant/Award Number: #297258; EPSRC, Grant/Award Number: EP/T015926/1

**Abstract**

We show that Mirzakhani’s curve counting theorem also holds if we replace surfaces by orbifolds.

**MSC 2020**

57K20, 57R18 (primary)

## 1 | INTRODUCTION

Throughout this paper, we let  $\Gamma \subset \mathrm{PSL}_2 \mathbb{R}$  be a non-elementary finitely generated discrete subgroup of the group of orientation preserving isometries of the hyperbolic plane  $\mathbb{H}^2$ . Suppose for a moment that  $\Gamma$  is torsion free, let  $S = \mathbb{H}^2/\Gamma$  be the associated hyperbolic surface, and let  $S(S)$  be the set of free homotopy classes of closed unoriented primitive essential curves therein. Here essential just means that the given homotopy class of curves is neither trivial nor peripheral. The mapping class group  $\mathrm{Map}(S)$  of  $S$  acts on  $S(S)$  and we say that two elements in the same orbit are *of the same type*. Mirzakhani studied the asymptotic behavior, when  $L \rightarrow \infty$ , of the number of elements in  $S(S)$  of some fixed type  $\gamma_0$  and with at most length  $L$ . More concretely, she proved in [13–15] that the limit

$$\lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} |\{\gamma \text{ of type } \gamma_0 \text{ with } \ell_S(\gamma) \leq L\}| \tag{1.1}$$

exists and is positive for every  $\gamma_0 \in S(S)$ . Here  $g$  is the genus of  $S$ ,  $r$  is the number of ends, and  $\ell_S(\gamma)$  is the length of the hyperbolic geodesic in the homotopy class  $\gamma$ .

© 2022 The Authors. *Transactions of the London Mathematical Society* is copyright © London Mathematical Society. This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

The goal of this note is to prove that this statement remains true when  $\Gamma$  has torsion, that is, when  $O = \mathbb{H}^2/\Gamma$  is an orbifold instead of a surface.

**Theorem 1.1.** *Let  $\Gamma \subset \text{PSL}_2 \mathbb{R}$  be a non-elementary finitely generated discrete subgroup and  $O = \mathbb{H}^2/\Gamma$  the associated 2-dimensional hyperbolic orbifold. Then the limit*

$$\lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} |\{\gamma \text{ of type } \gamma_0 \text{ with } \ell_O(\gamma) \leq L\}|$$

*exists and is positive for any  $\gamma_0 \in S^{\text{or}}(O)$ . Here  $g$  is the genus of the orbifold  $O$  and  $r$  is the sum of the numbers of singular points and ends.*

A few comments on the notation and terminology used in Theorem 1.1:

- (1) The topological space underlying an orientable 2-dimensional hyperbolic orbifold is an orientable topological surface. The genus of the orbifold is by definition the genus of that surface.
- (2) In the theorem, and also in the remaining of the paper,  $S^{\text{or}}(O)$  is the set of free homotopy classes of closed unoriented primitive essential curves, where the homotopy is taken in the category of orbifolds and where essential means that the curves in the given homotopy class are neither peripheral nor represent finite-order elements in the orbifold fundamental group  $\pi_1^{\text{or}}(O)$ . Accordingly,  $\ell_O(\gamma)$  is the length of the shortest curve homotopic to  $\gamma$  in the category of orbifolds. See Section 2.3 for details.
- (3) As for surfaces, two elements in  $S^{\text{or}}(O)$  are of the same type if they differ by an element of the mapping class group

$$\text{Map}^{\text{or}}(O) = \text{Homeo}^{\text{or}}(O)/\text{Homeo}_0^{\text{or}}(O).$$

Here,  $\text{Homeo}^{\text{or}}(O)$  is the group of homeomorphisms of  $O$  in the category of orbifolds, and  $\text{Homeo}_0^{\text{or}}(O)$  is its identity component. The mapping class group  $\text{Map}^{\text{or}}(O)$  is infinite unless  $O$  is *exceptional*, by what we mean that it has genus  $g = 0$  and that  $r = 3$ . See Section 2.6 for more details on the mapping class group of an orbifold.

As was already the case for the proof that we gave in [7] of Mirzakhani’s (1.1), we will derive Theorem 1.1 from the weak- $*$ -convergence of certain measures on the space  $C^{\text{or}}(O)$  of currents, which is the space of  $\pi_1^{\text{or}}(O)$ -invariant Radon measures on the set of geodesics on the orbifold universal cover  $\tilde{O}$  of  $O$ . Trusting that the reader is familiar with currents, we just recall at this point that the set  $\mathbb{R}_{\geq 0} S^{\text{or}}(O)$  of weighted curves is a dense subset of  $C^{\text{or}}(O)$ , that  $C^{\text{or}}(O)$  is a cone in a linear space, and that the action of  $\text{Map}^{\text{or}}(O)$  on  $S^{\text{or}}(O)$  extends to a linear action on  $C^{\text{or}}(O)$ . We will recall a few facts about currents in Section 4.1 below, but we do already at this point refer the reader to [1–4, 7] for details and background.

**Theorem 1.2.** *Let  $O$  be a compact orientable non-exceptional hyperbolic orbifold with possibly empty totally geodesic boundary and let  $C^{\text{or}}(O)$  be the associated space of geodesic currents. There is a Radon measure  $\mathbf{m}_{\text{Thu}}$  on  $C^{\text{or}}(O)$  such that for any  $\gamma_0 \in S^{\text{or}}(O)$ , we have*

$$\lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} \sum_{\gamma \in \text{Map}^{\text{or}}(O) \cdot \gamma_0} \delta_{\frac{1}{L}\gamma} = C(\gamma_0) \cdot \mathbf{m}_{\text{Thu}}$$

for some positive constant  $C(\gamma_0) > 0$ . Here  $g$  is the genus of the orbifold  $O$ ,  $r$  is the sum of the numbers of singular points and boundary components, and  $O$  is non-exceptional if  $(g, r) \neq (0, 3)$ . Moreover  $\delta_{\frac{1}{L}\gamma}$  stands for the Dirac measure on  $C^{\text{or}}(O)$  centered at  $\frac{1}{L}\gamma$ , and the convergence takes place with respect to the weak- $*$ -topology on the space of Radon measures on  $C^{\text{or}}(O)$ .

Again a few comments:

- (1) Theorem 1.2 remains true if we replace curves by multicurves, that is, if we replace  $\gamma_0$  by finite formal linear combinations (with positive coefficients) of elements in  $S^{\text{or}}(O)$ . In fact, the proof is just the same, only needing a bit more of notation to keep track of things, and the interested reader will have no difficulties making the necessary tweaks.
- (2) Also, as is the case for surfaces, the statement of Theorem 1.2 remains true if we replace  $\text{Map}^{\text{or}}(O)$  by a finite index subgroup  $G$ , and the constant on the right side changes exactly as it does in the case of surfaces — see [7, Exercise 8.2]. In fact, in the course of the proof of Theorem 1.2, we will have to work with such a finite index subgroup, the pure mapping class group.
- (3) The measure  $\mathfrak{m}_{\text{Thu}}$  in the statement of Theorem 1.2 arises as the push-forward under a certain map of the usual Thurston measure on the space of measured laminations of a surface. We will, however, also give a short intrinsic description of  $\mathfrak{m}_{\text{Thu}}$  in Section 4.4.
- (4) If one were to drop the assumption in Theorem 1.2 that the orbifold is non-exceptional, then the limit would trivially exist because the mapping class group would be finite, but the measure class of the obtained measure would obviously depend on  $\gamma_0$ . This is why we do need this assumption in Theorem 1.2 but not in Theorem 1.1 above or in Theorem 1.3 below.

All of this is nice and well and cute, but a more substantial observation is that Theorem 1.2 implies that Theorem 1.1 also holds if we replace  $\ell_O$  by many other notions of length: length with respect to a variable curvature metric, word length, extremal length, and so on. In fact, we can replace  $\ell_O$  by any continuous homogenous function

$$F : C^{\text{or}}(O) \rightarrow \mathbb{R}_{\geq 0}$$

on the space of currents, where homogeneous means that  $F(t \cdot \lambda) = t \cdot F(\lambda)$ . See [5, 12] for many examples of such functions.

**Theorem 1.3.** *Let  $O$  be a compact orientable hyperbolic orbifold with possibly empty totally geodesic boundary and let  $C^{\text{or}}(O)$  be the associated space of geodesic currents. Then the limit*

$$\lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} |\{\gamma \text{ of type } \gamma_0 \text{ with } F(\gamma) \leq L\}|$$

*exists and is positive for any  $\gamma_0 \in S^{\text{or}}(O)$  and any positive, homogenous, continuous function  $F : C^{\text{or}}(O) \rightarrow \mathbb{R}_{\geq 0}$ . Here  $g$  is the genus of the orbifold  $O$ , and  $r$  is the sum of the numbers of singular points and boundary components.*

Let us now describe the strategy of the proof of our main result, Theorem 1.2. Instead of aiming to give a stand-alone proof of the theorem along the lines of the proof in [7] of the corresponding result for surfaces, we are going to use the latter to obtain that for orbifolds. Suppose for the sake of

concreteness that  $O$  has no boundary and a single cone point and let  $\Sigma$  be the surface obtained by deleting from  $O$  a small ball around that singular point. The inclusion  $\Sigma \hookrightarrow O$  induces a surjective map

$$S(\Sigma) \rightarrow S(O) \cup \{*\}, \tag{1.2}$$

where  $*$  is just a point where one maps all essential curves in  $\Sigma$  which are non-essential in  $O$ . In fact, this map is equivariant under the isomorphism  $\text{Map}(\Sigma) \simeq \text{Map}^{\text{or}}(O)$  between the corresponding mapping class groups. Equivariance under this isomorphism implies that whenever  $\eta_0 \in S(\Sigma)$  maps to  $\gamma_0 \in S^{\text{or}}(O)$ , then the push-forward under (1.2) of the measure  $\frac{1}{L^{6g-6+2r}} \sum_{\eta \in \text{Map}(S) \cdot \eta_0} \delta_{\frac{1}{L}\eta}$  is the measure  $\frac{1}{L^{6g-6+2r}} \sum_{\gamma \in \text{Map}^{\text{or}}(O) \cdot \gamma_0} \delta_{\frac{1}{L}\gamma}$ . From the analog of Theorem 1.2 for surfaces, stated in Section 4.2 below, we get that the limit

$$\lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} \sum_{\eta \in \text{Map}(S) \cdot \eta_0} \delta_{\frac{1}{L}\eta}$$

exists. As we see, Theorem 1.2 would directly follow if the map (1.2) were to extend continuously to a map

$$C(\Sigma) \rightarrow C^{\text{or}}(O). \tag{1.3}$$

It is, however, easy to see that such an extension does not exist: for any three essential  $\alpha, \beta, \gamma \in \pi_1(\Sigma)$  with  $\beta \in \text{Ker}(\pi_1(\Sigma) \rightarrow \pi_1^{\text{or}}(O))$ , we have that  $\frac{1}{2n}[\alpha^n, \beta]\gamma$  converges when  $n \rightarrow \infty$  in  $C(\Sigma)$  to  $\alpha$  but is mapped to  $\frac{1}{2n}\gamma \in C^{\text{or}}(O)$  which converges to 0. We bypass this problem by choosing the representative  $\eta_0$  of  $\gamma_0$  so that the currents of the form  $\frac{1}{L}\eta$  with  $\eta \in \text{Map}(\Sigma) \cdot \eta_0$  are all contained in a closed subset of the set of currents on  $C(\Sigma)$  to which the map (1.2) actually extends continuously. We choose  $\eta_0$  to be *as simple as possible* in some precise sense given in Section 3. That the so-chosen  $\eta_0$  has the desired property follows from Proposition 3.3, the technical result at the core of this paper. This proposition basically asserts that the images in  $\tilde{O}$  of geodesics in  $\tilde{\Sigma}$  which are as simple as possible are uniformly quasigeodesic.

### 1.1 | Non-orientable orbifolds

It is known that, at least as stated, the limit (1.1) does not hold for non-orientable surfaces [9, 11] and this is why we assumed in the theorems above that the orbifold is orientable. It is, however, worth noting that all results here remain true for non-orientable orbifolds whose underlying topological space is an orientable surface. An example is  $D\Sigma/\tau$  where  $D\Sigma$  is the double of  $\Sigma$ , an orientable surface with boundary, and  $\tau$  is the involution interchanging the copies of the surface. The reason why the theorems remain true is that, up to passing to finite index subgroups, the mapping class group of such an orbifold is isomorphic to the mapping class group of an orientable surface for which we know that the analog of Theorem 1.2 holds. Anyways, we decided against extending the theorems above to this kind of non-orientable orbifolds because (1) it would make the paper much harder to read and (2) we do not have any concrete applications in mind.

## 1.2 | Plan of the paper

In Section 2 we recall some facts and definitions about orbifolds, maps between orbifolds, the mapping class group of orbifolds, and such. In Section 3 we state precisely what we mean by *as simple as possible* and state, without proof, Proposition 3.3. In Section 4 we recall a few facts about currents and, assuming Proposition 3.3, prove Theorem 1.2 and the other results mentioned above. In Section 5 we prove a few facts needed in Section 6, where we prove Proposition 3.3.

## 2 | ORBIFOLDS

In this section, we recall a few basics about orbifolds such as definitions, (hyperbolic) orbifolds as orbit spaces, and mapping class groups. We also fix some notation that we will use throughout the paper. This is why we also encourage readers who already know all about orbifolds to at least skim over this section.

### 2.1 | Orbifolds per se

An orbifold  $O$  is a space which is locally modeled on the quotient space of euclidean space by a finite group action. More precisely, an *orbifold chart* of a Hausdorff paracompact topological space  $O$  is a tuple  $(U, \hat{U}, \Gamma, \phi)$  where  $U \subset O$  and  $\hat{U} \subset \mathbb{R}^n$  are open, where  $\Gamma$  is a finite group acting on  $\hat{U}$ , and where  $\phi : \hat{U}/\Gamma \rightarrow U$  is a homeomorphism. An *orbifold atlas* is a collection  $\{(U_i, \hat{U}_i, \Gamma_i, \phi_i) \mid i \in I\}$  of orbifold charts such that  $\{U_i \mid i \in I\}$  is an open cover of  $O$  closed under intersections and such that whenever  $U_i \subset U_j$  there are (1) a group homomorphism  $f_{i,j} : \Gamma_i \rightarrow \Gamma_j$  and (2) an  $f_{i,j}$ -equivariant embedding

$$\hat{\phi}_{i,j} : \hat{U}_i \rightarrow \hat{U}_j$$

such that the diagram

$$\begin{array}{ccc} \hat{U}_i & \xrightarrow{\hat{\phi}_{i,j}} & \hat{U}_j \\ \downarrow & & \downarrow \\ \hat{U}_i/\Gamma_i & \longrightarrow & \hat{U}_j/\Gamma_j \\ \phi_i \downarrow & & \downarrow \phi_j \\ U_i & \xrightarrow{\subset} & U_j \end{array}$$

commutes. An *orbifold* is then a Hausdorff paracompact space endowed with an orbifold atlas.

The orbifold is orientable if all group actions  $\Gamma_i \curvearrowright \hat{U}_i$  and all embeddings  $\hat{\phi}_{i,j}$  are orientation preserving. Similarly if we replace orientable by smooth. An orbifold with boundary is defined in the same way but this time the sets  $\hat{U}_i$  are assumed to be open in  $\mathbb{R}_{\leq 0} \times \mathbb{R}^{n-1}$ . An  $n$ -dimensional hyperbolic orbifold is one where the sets  $\hat{U}_i$  are contained in  $\mathbb{H}^n$ , where the actions  $\Gamma_i \curvearrowright \hat{U}_i$  preserve the hyperbolic metric, and where the maps  $\hat{\phi}_{i,j}$  are isometric embeddings. To define what

is a hyperbolic orbifold with totally geodesic boundary, then one copies what we just wrote, only replacing  $\mathbb{H}^n$  by a closed half-space therein.

As is the case in the world of manifolds, orbifolds have maximal orbifold atlases, orientable orbifolds have maximal orientable orbifold atlases, smooth orientable orbifolds have maximal smooth orientable orbifold atlases, and so on. We will always assume that our orbifolds (with adjectives) are equipped with maximal atlases (with adjectives).

We refer to [18] for more on orbifolds.

### 2.2 | Singular points

A point  $p$  in an orbifold  $O$  is *singular* if there are an orbifold chart  $(U, \hat{U}, \Gamma, \phi)$  and  $\hat{p} \in \hat{U}$  with  $\phi(\hat{p}) = p$  and satisfying that  $\text{Stab}_\Gamma(\hat{p}) \neq \text{Id}$ . A point which is not singular is *regular*. We denote by  $\text{sing}(O)$  the set of singular points of  $O$ .

The *singular set*  $\text{sing}(O)$  is a closed subset of  $O$ . It might be empty, but also its complement might be empty. It is actually sometimes really important to allow oneself to work with orbifolds with  $\text{sing}(O) = O$  — not the simplest example one can find, but the moduli space  $\mathcal{M}_{2,0}$  of closed Riemann surfaces of genus 2 is such an orbifold. However,

all orbifolds in this paper are such that  $\text{sing}(O)$  is a proper subset of  $O$ .

In the cases we are interested in, namely, compact orbifolds which are orientable, connected, and 2-dimensional, we have that  $\text{sing}(O)$  is, in fact, a finite set of points in the interior of  $O$ .

*Remark.* Whenever we need to choose a base point in our orbifold  $O$ , for example, when working with the fundamental group  $\pi_1^{\text{or}}(O)$ , then we will assume without further mention that the base point is regular. The reader might amuse themselves by thinking about what the right notion of base point in the category of orbifolds would be if they allowed singular points to be base points.

### 2.3 | Maps between orbifolds

A map  $f : O \rightarrow O'$  between two orbifolds is then a continuous map such that whenever  $\phi_i : \hat{U}_i/\Gamma_i \rightarrow U_i$  and  $\phi'_i : \hat{U}'_i/\Gamma'_i \rightarrow U'_i$  are orbifolds charts for  $O$  and  $O'$  with  $f(U_i) \subset U'_i$ , then there are a homomorphism  $\Gamma_i \rightarrow \Gamma'_i$  and an equivariant continuous map  $\hat{f}_i : \hat{U}_i \rightarrow \hat{U}'_i$  such that the obvious diagram

$$\begin{array}{ccc} \hat{U}_i & \xrightarrow{\hat{f}_i} & \hat{U}'_i \\ \downarrow & & \downarrow \\ U_i & \xrightarrow{f} & U'_i \end{array}$$

commutes. If  $O$  and  $O'$  are smooth orbifolds and if the  $\hat{f}_i$  are smooth, then  $f$  is said to be smooth. Orbifolds and maps between orbifolds form a category. And the same for smooth orbifolds and smooth maps between them.

If  $(U, \hat{U}, \Gamma, \phi)$  is an orbifold chart of an orbifold  $O$ , then  $([0, 1] \times U, [0, 1] \times \hat{U}, \Gamma, \text{Id} \times \phi)$  where  $g \in \Gamma$  acts on  $[0, 1] \times \hat{U}$  via  $g(t, x) = (t, gx)$  is an orbifold chart of  $[0, 1] \times O$ . The collection of all so-obtained orbifold charts forms an orbifold atlas, giving  $[0, 1] \times O$  the structure of an orbifold. It thus makes sense to say that two orbifold maps

$$f, f' : O \rightarrow O'$$

are *homotopic in the category of orbifolds* if there is an orbifold map

$$F : [0, 1] \times O \rightarrow O', \quad F(t, x) = F_t(x)$$

with  $F_0 = f$  and  $F_1 = f'$ .

Anyways, armed with the notion of homotopy of orbifold maps, one can define the *orbifold fundamental group*  $\pi_1^{\text{or}}(O)$  of  $O$  exactly as one does for the usual fundamental group, just replacing homotopies by orbifold homotopies. One should note that any two orbifold maps which are homotopic as orbifold maps are also homotopic as maps between topological spaces, but that the converse does not need to be true. In fact, there are plenty of orbifolds which are simply connected as topological spaces but whose orbifold fundamental group is non-trivial, meaning that there are orbifold maps  $\gamma : \mathbb{S}^1 \rightarrow O$  which, as orbifold maps, are not homotopic to constant maps. As is the case for manifolds, in the category of orbifolds, free homotopy classes of *curves*  $\gamma : \mathbb{S}^1 \rightarrow O$  correspond to conjugacy classes in the orbifold fundamental group. We make the following convention:

If  $O$  is a compact orbifold with boundary then we will say that a curve is *essential* if it is not freely homotopic into the boundary and if the associated free homotopy class is that of an infinite order element in the orbifold fundamental group. We will also denote by  $S^{\text{or}}(O)$  the set of all free homotopy classes of essential curves in  $O$ .

## 2.4 | Orbifolds as orbit spaces

Following word-by-word the usual construction of the universal cover of a manifold but replacing homotopies by homotopies of orbifold maps, one gets the orbifold universal cover  $\tilde{O}$  of the orbifold  $O$ . As is the case for manifolds, the fundamental group  $\pi_1^{\text{or}}(O)$  acts discretely on the universal cover  $\tilde{O}$ . Similarly, orbifold maps  $f : O \rightarrow O'$  between orbifolds induce homomorphisms  $f_* : \pi_1^{\text{or}}(O) \rightarrow \pi_1^{\text{or}}(O')$  between the associated orbifold fundamental groups and lift to  $f_*$ -equivariant maps  $\tilde{f} : \tilde{O} \rightarrow \tilde{O}'$  between the universal covers.

Orbifolds whose universal cover is a manifold are said to be *good*. And they deserve that name because working with them is much easier than working with general orbifolds. For example, there is a pretty concrete description of the orbifold charts for good orbifolds  $O$ . They are namely of the form

$$\phi : \hat{U}/H \rightarrow U,$$

where  $H \subset \pi_1^{\text{or}}(O)$  is a finite subgroup, where  $\hat{U} \subset \tilde{O}$  is an open connected subset with  $H\hat{U} = \hat{U}$  and  $g\hat{U} \cap \hat{U} = \emptyset$  whenever  $g \notin H$ , where  $U$  is the image of  $\hat{U}$  under the universal covering map  $\pi : \tilde{O} \rightarrow O$ , and where finally  $\phi$  is the map given by  $\phi(xH) = \pi(x)$ .

Hyperbolic orbifolds, if one wants with geodesic boundary, are good. These are the orbifolds we will be interested in. We fix now the notation that we will be using from this point on:

**Notation.** Let  $\tilde{O} \subset \mathbb{H}^2$  be a closed connected (2-dimensional) subset of the hyperbolic plane with possibly empty geodesic boundary, let  $\Gamma \subset \text{PSL}_2 \mathbb{R}$  be a discrete subgroup which preserves  $\tilde{O}$  and such that the induced action  $\Gamma \curvearrowright \tilde{O}$  is cocompact, and denote by

$$O = \tilde{O}/\Gamma$$

the associated hyperbolic orbifold. When needed, we will refer to the hyperbolic metric on both  $O$  and  $\tilde{O}$  by  $\rho_{\text{hyp}}$ . Finally, we also write

$$\text{sing}(\Gamma) = \{p \in \tilde{O} \text{ with } \text{Stab}_\Gamma(p) \neq \text{Id}\}$$

for the set of points in  $\tilde{O}$  with non-trivial stabilizer, that is the preimage of  $\text{sing}(O)$  under the map  $\tilde{O} \rightarrow O$ .

In this setting,  $\tilde{O}$  is the orbifold universal cover of  $O$  and  $\Gamma = \pi_1^{\text{or}}(O)$  is its orbifold fundamental group.

It is not hard to see that the orbifolds  $O$  we are interested in are homeomorphic as topological spaces to surfaces, that is, to 2-dimensional manifolds. Such homeomorphisms do, however, destroy the orbifold structure. In fact, much more information is encoded in the surface that we get by deleting the singular points of  $O$ . Since we want to work with compact surfaces, we instead delete small balls around the singular points.

### 2.5 | The surface associated to a hyperbolic orbifold $O$

Continuing with the same notation, let  $O = \tilde{O}/\Gamma$  be a compact orientable hyperbolic 2-orbifold with possibly empty totally geodesic boundary. We choose now two positive constants  $\epsilon$  and  $\delta$  which will accompany us throughout the paper. Other than being very small, say  $\epsilon < 10^{-10}$ , here are the conditions that  $\epsilon$  has to satisfy:

- (C1)  $200\epsilon$  is less than the length of the shortest non-trivial periodic  $\rho_{\text{hyp}}$ -geodesic in  $O$ ,
- (C2)  $200\epsilon$  is less than the minimal distance between any two points in  $\text{sing}(\Gamma)$ , and
- (C3)  $200\epsilon$  is less than the distance between any point in  $\partial\tilde{O}$  and any point in  $\text{sing}(\Gamma)$ .

When it comes to  $\delta$ , we will later give a fourth condition (see (C4) in Section 5) that it has to satisfy but for now we just assume that  $3\delta < \epsilon$ . Note that this implies that the  $\delta$ -balls around points in  $\text{sing}(\Gamma)$  are disjoint of each other and do not meet  $\partial\tilde{O}$ . This means that

$$\hat{\Sigma} = \tilde{O} \setminus \mathcal{N}_{\text{hyp}}(\text{sing}(\Gamma), \delta) \tag{2.1}$$

is a smooth surface with boundary, where

$$\mathcal{N}_{\text{hyp}}(X, r) = \{p \in \tilde{O} \text{ with } d_{\text{hyp}}(p, X) < r\}.$$



Note that the action of  $\Gamma = \pi_1^{\text{or}}(\mathcal{O})$  on  $\tilde{\mathcal{O}}$  induces an action on  $\hat{\Sigma}$  which is not only discrete but also free. We refer to the quotient surface

$$\Sigma = \hat{\Sigma}/\Gamma$$

as *the surface associated to the orbifold*  $\mathcal{O}$  and denote its universal cover by  $\tilde{\Sigma}$ . By construction, it is also the universal cover of  $\hat{\Sigma}$ . In fact,  $\tilde{\Sigma}$  is the cover of  $\Sigma$  corresponding to the normal subgroup of  $\pi_1(\Sigma)$  generated by all loops homotopic into  $\partial\Sigma \setminus \partial\mathcal{O}$ .

*Remark.* We denote by  $B_{\text{hyp}}(q, r) \subset \tilde{\mathcal{O}}$  the hyperbolic ball of radius  $r$  around  $q$ . Equivalently,

$$B_{\text{hyp}}(q, r) = \mathcal{N}_{\text{hyp}}(\{q\}, r).$$

Also, abusing terminology, we will not distinguish between  $\mathcal{N}_{\text{hyp}}(X, r)$  or  $B_{\text{hyp}}(q, r)$  and their closures. That is, both open balls and closed balls, and open neighborhoods and closed neighborhoods are denoted using the same symbol.

## 2.6 | Mapping class groups of the orbifold and of the associated surface

As is the case for manifolds, one can say anything one wants to say about the orbifold  $\mathcal{O} = \tilde{\mathcal{O}}/\Gamma$  in terms of  $\Gamma$ -equivariant objects in the universal cover  $\tilde{\mathcal{O}}$ . For example, the group  $\text{Homeo}^{\text{or}}(\mathcal{O})$  of orbifold self-homeomorphisms of  $\mathcal{O}$  can be identified with

$$\text{Homeo}^{\text{or}}(\mathcal{O}) = \text{Homeo}_{\Gamma}(\tilde{\mathcal{O}})/\Gamma,$$

where

$$\text{Homeo}_{\Gamma}(\tilde{\mathcal{O}}) = \{\tilde{f} \in \text{Homeo}(\tilde{\mathcal{O}}) \text{ with } \tilde{f}\Gamma\tilde{f}^{-1} = \Gamma\}$$

is the group of (topological) homeomorphisms of  $\tilde{\mathcal{O}}$  conjugating  $\Gamma$  to itself. The mapping class group, in the category of orbifolds, of  $\mathcal{O}$  is then the group

$$\text{Map}^{\text{or}}(\mathcal{O}) = \text{Homeo}^{\text{or}}(\mathcal{O})/\text{Homeo}_0^{\text{or}}(\mathcal{O}),$$

where  $\text{Homeo}_0^{\text{or}}(\mathcal{O})$  is the identity component of  $\text{Homeo}^{\text{or}}(\mathcal{O})$ .

The group  $\text{Homeo}_{\Gamma}(\tilde{\mathcal{O}})$  acts on the set

$$\text{sing}(\Gamma) = \{p \in \tilde{\mathcal{O}} \text{ with } \text{Stab}_{\Gamma}(p) \neq \text{Id}\}$$

of points with non-trivial stabilizer. It also acts on the set  $\pi_0(\partial\tilde{\mathcal{O}})$  of boundary component of  $\tilde{\mathcal{O}}$ . It follows that the mapping class group acts on the finite sets  $\text{sing}(\Gamma)/\Gamma$  and  $\pi_0(\partial\tilde{\mathcal{O}})/\Gamma$ . The *pure mapping class group*

$$\text{PMap}^{\text{or}}(\mathcal{O}) = \{\phi \in \text{Map}^{\text{or}}(\mathcal{O}) \text{ pointwise fixing } \text{sing}(\Gamma)/\Gamma \text{ and } \pi_0(\partial\tilde{\mathcal{O}})/\Gamma\}$$

is the finite index subgroup of  $\text{Map}^{\text{or}}(\mathcal{O})$  consisting of mapping classes which act trivially on these two sets.

Note now that the canonical inclusion  $\Sigma \hookrightarrow \mathcal{O}$  into our orbifold of the associated surface is an embedding in the category of orbifolds. We have, however, also other interesting maps  $\Sigma \rightarrow \mathcal{O}$ , namely, those which are the identity outside of a small neighborhood of the boundary of  $\Sigma$  and which map  $\Sigma \setminus \partial\Sigma$  homeomorphically to  $\mathcal{O} \setminus \text{sing}(\mathcal{O})$ . Such maps are not homeomorphisms but they induce homomorphisms between the group of homeomorphisms of  $\Sigma$  acting trivially on  $\pi_0(\partial\Sigma)$  and the group of orbifold homeomorphisms of  $\mathcal{O}$ . Any such map induces an isomorphism between the pure mapping class groups

$$\text{PMap}(\Sigma) \simeq \text{PMap}^{\text{or}}(\mathcal{O}) \tag{2.2}$$

of  $\Sigma$  and  $\mathcal{O}$ , where

$$\text{PMap}(\Sigma) = \{\phi \in \text{Homeo}(\Sigma) \text{ acting trivially on } \pi_0(\partial\Sigma)\} / \text{Homeo}_0(\Sigma).$$

It is well known that every mapping class in  $\text{Map}(\Sigma)$  can be represented by a diffeomorphism.

Although our definition of the mapping class group differs from theirs (we do not have twists around the boundary), we refer to the book [8] by Farb and Margalit for background on the mapping class group.

### 2.7 | A metric on the associated surface $\hat{\Sigma}$

Although (locally) negatively curved from the point of view of comparison geometry, the restriction of the hyperbolic metric  $\rho_{\text{hyp}}$  to  $\hat{\Sigma}$  is not as nice as one would wish. The problem is that, since the new boundary components are concave, geodesics are not uniquely determined by their tangent vectors at a point. In particular, distinct geodesics do not need to be transversal to each other. This is why we from now on endow  $\hat{\Sigma}$  with a smooth Riemannian metric  $\rho$  with the following properties.

- $\rho$  is negatively curved and  $\Gamma$ -invariant.
- The boundary of  $\hat{\Sigma}$  is totally geodesic with respect to  $\rho$ .
- Both  $\rho$  and  $\rho_{\text{hyp}}$  agree on the subset  $\tilde{\mathcal{O}} \setminus \mathcal{N}_{\text{hyp}}(\text{sing}(\Gamma), 2\delta)$  of  $\hat{\Sigma}$ .
- If  $I \subset \tilde{\mathcal{O}}$  is a  $\rho_{\text{hyp}}$ -geodesic segment starting at a point  $\text{sing}(\Gamma)$  and with  $\rho_{\text{hyp}}$ -length  $3\delta$ , then  $I \cap \hat{\Sigma}$  is a  $\rho$ -geodesic segment perpendicular to the boundary of  $\Sigma$ .

The reason why we impose this final condition is that if  $p \in \text{sing}(\Gamma)$  and  $r > 2\delta$  are such that  $\rho$  and  $\rho_{\text{hyp}}$  agree on  $B_{\text{hyp}}(p, r) \setminus B_{\text{hyp}}(p, 2\delta)$ , then the  $\rho_{\text{hyp}}$  radial foliation  $\mathcal{F}$  of  $B_{\text{hyp}}(p, r) \setminus B_{\text{hyp}}(p, \delta)$  is  $\rho$ -geodesic. It follows in particular that the restriction of the radial projection

$$B_{\text{hyp}}(p, r) \setminus B_{\text{hyp}}(p, \delta) \rightarrow \partial B_{\text{hyp}}(p, r)$$

to any  $\rho$ -geodesic segment  $\eta$  which is not contained in a leaf of  $\mathcal{F}$  is monotonic in the sense that its derivative is never 0. We thus get that simple  $\rho$ -geodesic segments  $\eta \subset B_{\text{hyp}}(p, r) \setminus B_{\text{hyp}}(p, \delta)$  whose endpoints are in  $\partial B_{\text{hyp}}(p, r)$  and meet each leaf of  $\mathcal{F}$  at most once — see Figure 1. We record this fact for later use:

**Lemma 2.1.** *Suppose that  $p \in \text{sing}(\Gamma)$  and  $r > 2\delta$  are such that  $\rho$  and  $\rho_{\text{hyp}}$  agree on  $B_{\text{hyp}}(p, r) \setminus B_{\text{hyp}}(p, 2\delta)$ , and let  $\eta \subset B_{\text{hyp}}(p, r) \setminus B_{\text{hyp}}(p, \delta)$  be a  $\rho$ -geodesic segment whose boundary points are*

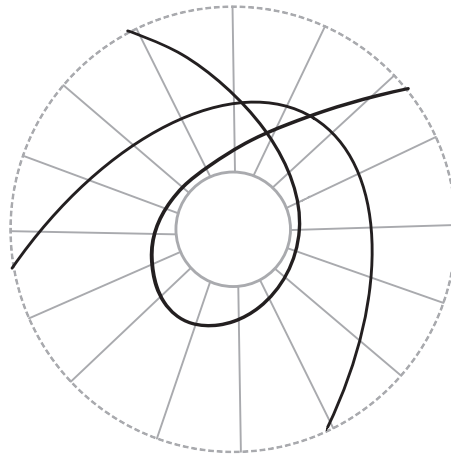


FIGURE 1 Schematic representation of two  $\rho$ -geodesic segments in  $B_{\text{hyp}}(p, 3\delta) \setminus B_{\text{hyp}}(p, \delta)$

contained in  $\partial B_{\text{hyp}}(p, r)$ . If  $\eta$  is simple, then  $\eta$  meets every  $\rho_{\text{hyp}}$ -geodesic ray emanating out of  $p$  at most once. In particular,  $\eta$  has at most length  $2\pi \sinh(r)$ .

We should comment on the existence of  $\rho$ . In fact, it is not hard to construct such a metric. For example, when working in standard hyperbolic polar coordinates  $(r, \theta)$  in the ball  $B_{\text{hyp}}(p, 3\delta)$  around  $p \in \text{sing}(\Gamma)$ , one can take any

$$\rho = dr^2 + \phi(r)^2 \cdot d\theta^2, \tag{2.3}$$

where  $\phi : [\delta, 3\delta) \rightarrow (0, \infty)$  is a smooth function satisfying

$$\phi''(\cdot) > 0, \phi'(\delta) = 0, \text{ and } \phi(s) = \sinh(s) \text{ for } s > 2\delta.$$

The first condition on  $\phi$  ensures that the sectional curvature  $\kappa = \frac{-\phi''}{\phi}$  is negative, the second that  $\{d_{\text{hyp}}(p, \cdot) = \delta\}$  is totally geodesic, and the third that  $\rho$  agrees with  $\rho_{\text{hyp}}$  on  $B_{\text{hyp}}(p, 3\delta) \setminus B_{\text{hyp}}(p, 2\delta)$ . In particular, if we use the same function  $\phi$  on each  $3\delta$ -ball around points in  $\text{sing}(\Gamma)$  and we set  $\rho = \rho_{\text{hyp}}$  outside those balls, then we obtain a  $\Gamma$ -invariant metric on the whole of  $\hat{\Sigma}$ . Finally, note that the curves  $t \mapsto (t, \theta)$ , that is, the  $\rho_{\text{hyp}}$ -geodesic segments starting at  $p$ , are  $\rho$ -geodesic segments for any choice of  $\phi$ . In other words, also the fourth property we wanted our metric to satisfy holds.

Note that  $\Gamma$ -invariance of  $\rho$  implies that it descends to a metric on  $\Sigma$  which we once again call  $\rho$ . Similarly, we denote also by  $\rho$  the induced metric on the universal cover  $\tilde{\Sigma}$ .

### 3 | AS SIMPLE AS POSSIBLE REPRESENTATIVES

Continuing with the same notation, let  $O = \tilde{O}/\Gamma$  be a compact orientable hyperbolic orbifold and let  $\Sigma = \hat{\Sigma}/\Gamma$  with  $\hat{\Sigma}$  as in (2.1) be the associated surface, endowed with the metric  $\rho$  we just fixed. By construction,  $\hat{\Sigma}$  is a connected subset of the universal cover  $\tilde{O}$  of  $O$ . It follows that the inclusion

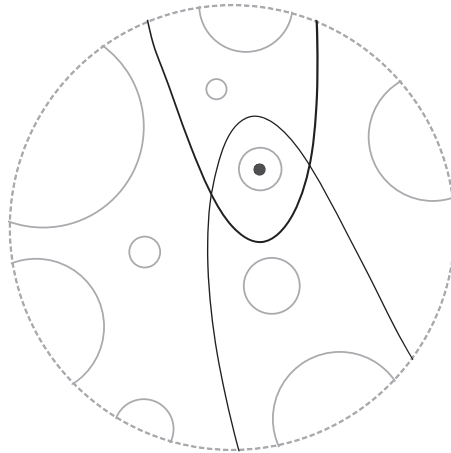


FIGURE 2 Two geodesics (in black) in  $\hat{\Sigma}$  that differ by an element in  $\Gamma$  fixing the marked point: they are not as simple as possible

$\Sigma \hookrightarrow \mathcal{O}$  induces a surjective homomorphism

$$\pi_1(\Sigma) \rightarrow \pi_1^{\text{or}}(\mathcal{O}) = \Gamma.$$

This means that every homotopically essential curve in  $\mathcal{O}$  is freely homotopic (in the category of orbifolds) to one contained in  $\Sigma$ . In this section, we describe how to pick for curves in  $\mathcal{O}$  representatives in  $\Sigma$  that are as simple as possible.

**Definition.** A  $\rho$ -geodesic  $\alpha : \mathbb{R} \rightarrow \hat{\Sigma}$  whose image is not contained in  $\partial\hat{\Sigma}$  is *as simple as possible* if

- (1) it is injective, and
- (2) for all  $g \in \Gamma$ , the geodesics  $\alpha$  and  $g(\alpha)$  are either identical or meet at most once.

We say that a  $\rho$ -geodesic in  $\Sigma$  is as simple as possible if its lifts to  $\hat{\Sigma}$  are as simple as possible. Similarly, a  $\rho$ -geodesic in the universal cover  $\hat{\Sigma}$  of  $\Sigma$  is as simple as possible if its images in  $\hat{\Sigma}$  are as simple as possible. Finally, a homotopy class in  $\Sigma$  is as simple as possible if its  $\rho$ -geodesic representative is as simple as possible (Figure 2).

Before going any further, we note that non-trivial closed  $\rho_{\text{hyp}}$ -geodesics  $\gamma : \mathbb{S}^1 \rightarrow \mathcal{O}$  have representatives  $\eta : \mathbb{S}^1 \rightarrow \Sigma \subset \mathcal{O}$  that are as simple as possible. It suffices to choose  $\eta \subset \Sigma$  to be a shortest representative of  $\gamma$ . Indeed, the fact that  $\eta$  is shortest implies that its lifts to  $\hat{\Sigma}$  have no bigons, showing that  $\eta$  is as simple as possible. We record this fact for later use:

**Lemma 3.1.** *Every  $\rho_{\text{hyp}}$ -geodesic  $\gamma : \mathbb{S}^1 \rightarrow \mathcal{O}$  is freely homotopic, in the category of orbifolds, to a  $\rho$ -geodesic  $\eta : \mathbb{S}^1 \rightarrow \Sigma \subset \mathcal{O}$  which is as simple as possible.*

The reader might be wondering why instead of simply speaking of shortest representatives we choose something as clumsy as “as simple as possible.” The reason is that the latter property is mapping class group invariant:

**Lemma 3.2.** *If  $\eta \subset \Sigma$  is as simple as possible, then  $\phi(\eta)$  is also as simple as possible for every  $\phi \in \text{PMap}(\Sigma)$ .*

*Proof.* Abusing notation, denote the  $\rho$ -geodesic freely homotopic to  $\eta$  by the same letter. To determine  $\phi(\eta)$  choose first a representative  $\varphi \in \text{Diff}(\Sigma)$  of the mapping class  $\phi$  and let  $\varphi_* : \pi_1(\Sigma) \rightarrow \pi_1(\Sigma)$  be the homomorphism induced by  $\varphi$  — we can always choose  $\varphi$  so that it fixes some point and take that point as the base point for the fundamental group. Note that  $\varphi_*$  preserves the normal subgroup of  $\pi_1(\Sigma)$  generated by loops freely homotopic into  $\partial\Sigma \setminus \partial O$  and that  $\pi_1(\hat{\Sigma}) \subset \pi_1(\Sigma)$  is nothing other than this subgroup. We get that  $\varphi$  lifts to  $\hat{\Sigma}$ , or more precisely, that there is a  $\varphi_*$ -equivariant lift  $\hat{\varphi} \in \text{Diff}(\hat{\Sigma})$ . Now, if  $\hat{\eta}$  is a lift of  $\eta$  to  $\hat{\Sigma}$ , then we have for all  $g \in \Gamma$  that

$$\hat{\varphi}(\hat{\eta}) \cap \varphi_*(g)(\hat{\varphi}(\hat{\eta})) = \hat{\varphi}(\hat{\eta}) \cap \hat{\varphi}(g\hat{\eta}) = \hat{\varphi}(\eta \cap g\eta).$$

Since  $\eta$  was as simple as possible, we get that  $\hat{\varphi}(\hat{\eta})$  is simple and that  $\hat{\varphi}(\hat{\eta}) \cap \varphi_*(g)\hat{\varphi}(\hat{\eta})$  intersects its individual  $\Gamma$ -translates at most once, and that if these intersections take place, then they are transversal to each other.

It follows that the image of  $\varphi(\eta)$  in  $\Sigma$  has no bigons. The same is true for  $(\phi(\eta))_*$ , the geodesic in  $(\Sigma, \rho)$  freely homotopic to  $\varphi(\eta)$ . Now, [10, Theorem 2.1] implies that these two curves are not only freely homotopic to each other but also transversely freely homotopic to each other. This means in particular that intersection points are neither destroyed nor created during the homotopy. Hence, each lift of  $(\phi(\eta))_*$  to  $\hat{\Sigma}$  meets its individual  $\Gamma$ -translates in at most one point. In other words,  $\phi(\eta)$  is as simple as possible.  $\square$

The reason why we are interested in  $\rho$ -geodesics in  $\hat{\Sigma}$  which are as simple as possible is that, as we will see shortly, this topological property implies that they are uniform quasigeodesics with respect to the hyperbolic metric. Recall that a continuous curve  $\alpha : \mathbb{R} \rightarrow \tilde{O}$  is *A-quasigeodesic* if we have

$$A \cdot |t - s| + A \geq d_{\text{hyp}}(\alpha(s), \alpha(t)) \geq \frac{1}{A}|s - t| - A$$

for all  $s, t \in \mathbb{R}$ . It is quasigeodesic if it is  $A$ -quasigeodesic for some  $A \geq 1$ .

We are now ready to state the key technical result of this paper.

**Proposition 3.3.** *Let  $O$  be as in the statement of Theorem 1.2,  $\hat{\Sigma}$  as in (2.1), and  $\rho$  the metric on  $\hat{\Sigma}$  constructed in Section 2.7. There exists  $A \geq 1$  such that any unit speed  $\rho$ -geodesic  $\alpha : \mathbb{R} \rightarrow \hat{\Sigma}$  which is (1) a quasigeodesic in  $(\tilde{O}, \rho_{\text{hyp}})$  and (2) as simple as possible is actually  $A$ -quasigeodesic in  $(\tilde{O}, \rho_{\text{hyp}})$ .*

Proposition 3.3 will be proved in Section 6. We just add now a few comments.

- (1) Note that in Proposition 3.3 we cannot simply drop the assumption that  $\alpha$  is a quasigeodesic in  $(\tilde{O}, \rho)$ . For example, if  $\alpha$  is a lift of a simple geodesic in  $\Sigma$  which spirals in both directions onto components of  $\partial\Sigma \setminus \partial O$ , then it is as simple as possible but not a quasigeodesic and in particular not an  $A$ -quasigeodesic for any choice of  $A$ . However, this is basically the only case we have to rule out because we could replace the condition that  $\alpha$  is a quasigeodesic in the proposition by the assumption that  $\alpha$  does not accumulate on a compact component of  $\partial\hat{\Sigma}$  in either direction. We leave it, however, as it is because the curves we will be interested in are

automatically quasigeodesics: they are lifts  $\hat{\eta}$  to  $\hat{\Sigma}$  of representatives  $\eta$  in  $\Sigma$  of essential curves  $\gamma$  in  $O$ .

- (2) Suppose that  $\gamma \in S^{\text{or}}(O)$  is an essential curve in  $O$ . Proposition 3.3 implies that we have

$$\ell_\rho(\eta) \leq A \cdot \ell_{\text{hyp}}(\gamma)$$

for any  $\rho$ -geodesic representative  $\eta \subset \Sigma$  of  $\gamma$  which is as simple as possible. It follows that  $\gamma$  only has finitely many such representatives.

- (3) On the other hand, as simple as possible representatives are not unique. In fact, a primitive closed geodesic in  $O$  which goes through  $k$  cone points of odd order and none of even order, has at least  $2^k$  representatives in  $\Sigma$  which are as simple as possible: at each one of those  $k$  points, steer slightly either right or left to avoid hitting the cone point. And this is not optimal because perturbing the metric slightly one can get the geodesic off  $\text{sing}(O)$ , and representatives that were as simple as possible stay as simple as possible.
- (4) The construction sketched in (3) shows that “shortest” and “as simple as possible” are not the same thing.

## 4 | MAIN RESULTS

In this section we prove Theorem 1.2 assuming Proposition 3.3. However, before doing so we have to recall a few facts about currents and about Mirzakhani’s counting theorem.

### 4.1 | Currents

Let  $X$  be a simply connected negatively curved surface with possibly empty totally geodesic boundary, and let  $G \subset \text{Isom}_+(X)$  be a discrete subgroup of orientation preserving isometries with  $X/G$  compact. We will be interested in the following two possible cases:

- $X = \tilde{O} \subset \mathbb{H}^2$  is the universal cover of our compact hyperbolic orbifold  $O = \tilde{O}/\Gamma$  and  $G = \pi_1^{\text{or}}(O) = \Gamma$  is its fundamental group.
- $X = (\tilde{\Sigma}, \rho)$  is the universal cover of  $\Sigma$  endowed with the metric  $\rho$  and  $G = \pi_1(\Sigma)$  is its fundamental group.

Let  $\mathcal{G}(X)$  be the set of all unoriented bi-infinite geodesics in  $X$ . The action  $G \curvearrowright X$  induces an action of  $G$  on  $\mathcal{G}(X)$ . A current on  $X/G$  is a  $G$ -invariant Radon measure on  $\mathcal{G}(X)$ . Let  $C^{\text{or}}(X/G)$  be the space of all currents on  $X/G$  endowed with the weak- $*$ -topology. It is a Hausdorff, metrizable, second countable, and locally compact space, and the projectivized space  $\mathbb{P}C^{\text{or}}(X/G)$  is compact.

*Remark.* We insist that  $X/G$ , and thus our orbifold  $O$ , is compact because this is what guarantees that the space  $C^{\text{or}}(X/G)$  is locally compact.

There are plenty of currents. In fact there is a natural homeomorphism between  $C^{\text{or}}(X/G)$  and the space of geodesic flow invariant Radon measures on the projectivized unit tangent bundle  $PT^1X/G$  supported by the set of bi-infinite orbits. For example, every primitive closed unit speed geodesic  $\gamma$  in  $\mathcal{G}(X)$ , or equivalently every unoriented periodic orbit of the geodesic flow, yields a

geodesic flow invariant measure on  $PT^1X/G$ : the measure of  $U$  is the arc length of  $\gamma \cap U$ . The current associated to this measure is called the *counting current associated to the geodesic  $\gamma$* . The counting current determines the original geodesic  $\gamma$  — this justifies referring to the current and the geodesic by the same letter — and the name is explained because, for a set of geodesics  $V \subset \mathcal{G}(X)$  the value of  $\gamma(V)$  is nothing other than the number of lifts of  $\gamma$  to  $X$  which belong to  $V$ .

Note that every essential curve  $\gamma$  in  $X/G$  (in the sense that we gave to the word essential at the end of Section 2.3) is freely homotopic to a unique geodesic  $\gamma_*$  in  $X/G$ . In this case we denote the associated counting current by  $\gamma$  instead of  $\gamma_*$ . We hope that this will not cause any confusion.

*Remark.* If the action of  $G$  on  $X$  is free, then we drop the superscript “or.” For example, we write  $C(\Sigma)$  instead of  $C^{\text{or}}(\Sigma)$ . We use this superscript to avoid mixing up currents for the orbifold  $X/G$  and currents for the underlying topological surface.

Currents were introduced by Bonahon and we refer to his papers [2–4] for details and background. See also [1]. However, although all these sources are highly recommended, the reader will not be surprised on hearing that we will mostly follow the same notation and terminology as in our book [7].

## 4.2 | Mirzakhani’s counting theorem

As we mentioned already in the introduction, Theorem 1.2 is well known in the case that we are working with surfaces instead of orbifolds. In that case we have the following result [7, Theorem 8.1].

**Mirzakhani’s counting theorem.** *Let  $\Sigma$  be a compact connected orientable surface of genus  $g$  and with  $r$  boundary components and suppose that  $3g - 3 + r > 0$ . Let also  $\eta_0 \subset \Sigma$  be a homotopically primitive essential curve. Then there are constants  $c^{\text{PMap}}(\eta_0), \mathfrak{b}_{g,r}^{\text{PMap}} > 0$  such that*

$$\lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} \sum_{\eta \in \text{PMap}(\Sigma) \cdot \eta_0} \delta_{\frac{1}{L}\eta} = \frac{c^{\text{PMap}}(\eta_0)}{\mathfrak{b}_{g,r}^{\text{PMap}}} \cdot \mathfrak{m}_{\text{Thu}}^{\Sigma}.$$

Here  $\delta_{\frac{1}{L}\eta}$  is the Dirac measure on  $C(\Sigma)$  centered at the current  $\frac{1}{L}\eta$ ,  $\mathfrak{m}_{\text{Thu}}^{\Sigma}$  is the Thurston measure on  $C(\Sigma)$ , and the convergence takes place with respect to the weak- $*$ -topology on the space of Radon measures on  $C(\Sigma)$ .

*Remark.* The counting theorem remains true if we replace the pure mapping class group by any other finite index subgroup of the mapping class group. However, the obtained multiple of the Thurston measure depends on the subgroup in question. This explains the superscript PMap in the constants  $c^{\text{PMap}}(\eta_0)$  and  $\mathfrak{b}_{g,r}^{\text{PMap}}$ . See [7, Exercise 8.2] for explicit formulas for the dependence of the constants on the chosen subgroup of the mapping class group.

Since we named the above theorem after Maryam Mirzakhani while referring to our book [7], we should add a brief comment on the genesis of this theorem. For simple curves, Mirzakhani proved this theorem in [13, 14] but for general curves the history is slightly more complicated. To

explain why, note that the theorem implies that

$$\lim_{L \rightarrow \infty} \frac{\{\gamma \in \text{PMap}(\Sigma) \text{ with } F(\gamma) \leq L\}}{L^{6g-6+2r}} = \frac{c^{\text{PMap}}(\gamma_0)}{\mathfrak{b}_{g,r}^{\text{PMap}}} \cdot m_{\text{Thu}}^\Sigma(\{F(\cdot) \leq 1\})$$

whenever  $F : \mathcal{C}(\Sigma) \rightarrow \mathbb{R}_+$  is a continuous, positive, and homogenous function (compare with [7, Theorem 9.1] or with the proof of Theorem 1.3 below). In [15] Maryam proved the existence of the latter limit in the case that  $F$  is the hyperbolic length. At the same time, we were also investigating the same problem and we proved in [6] that every sublimit of the sequence in the counting theorem has a subsequence which converges to a multiple of the Thurston measure. The existence of the limit in the counting theorem follows if one combines these two facts [6, 15]. This was clear to both Mirzakhani and ourselves at the time. A problem with that state of affairs was that Mirzakhani’s arguments and ours come from different places, and this made things a bit too opaque. For example, there was some confusion about the chosen normalization for the Thurston measure. This meant that it was not obvious how the arising constants should be understood (this problem was, to some extent, solved in [16, 17]). Finally, or maybe finally for the time being, a unified proof of the counting theorem as stated above was provided in [7]. The constants  $c^{\text{PMap}}(\gamma_0)$  and  $\mathfrak{b}_{g,r}^{\text{PMap}}$  in the statement of the counting theorem above are as given in [7, Chapter 8], and the Thurston measure is defined to be the scaling limit

$$m_{\text{Thu}}^\Sigma = \lim_{L \rightarrow \infty} \sum_{\gamma \in \mathcal{ML}_Z} \delta_{\frac{1}{L}\gamma}. \tag{4.1}$$

Anyways, let us return to the concrete topic of this paper.

### 4.3 | Proof of Theorem 1.2

We prove now our main theorem assuming Proposition 3.3. As mentioned earlier, the proposition will be proved in Section 6.

**Theorem 1.2.** *Let  $O$  be a compact orientable non-exceptional hyperbolic orbifold with possibly empty totally geodesic boundary and let  $C^{\text{or}}(O)$  be the associated space of geodesic currents. There is a Radon measure  $m_{\text{Thu}}$  on  $C^{\text{or}}(O)$  such that for any  $\gamma_0 \in S^{\text{or}}(O)$  we have*

$$\lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} \sum_{\gamma \in \text{Map}^{\text{or}}(O) \cdot \gamma_0} \delta_{\frac{1}{L}\gamma} = C(\gamma_0) \cdot m_{\text{Thu}}$$

for some positive constant  $C(\gamma_0) > 0$ . Here  $g$  is the genus of the orbifold  $O$ ,  $r$  is the sum of the numbers of singular points and boundary components, and  $O$  is non-exceptional if  $(g, r) \neq (0, 3)$ . Moreover  $\delta_{\frac{1}{L}\gamma}$  stands for the Dirac measure on  $C^{\text{or}}(O)$  centered at  $\frac{1}{L}\gamma$ , and the convergence takes place with respect to the weak- $*$ -topology on the space of Radon measures on  $C^{\text{or}}(O)$ .

As we already mentioned in the introduction, the idea of the proof is to show that our given homotopy class  $\gamma_0$  has a representative  $\eta_0$  in  $\Sigma$  such that the measures

$$\frac{1}{L^{6g-6+2r}} \sum_{\eta \in \text{PMap}(\Sigma) \cdot \eta_0} \delta_{\frac{1}{L}\eta}$$



inside the limit in Mirzakhani's counting theorem are supported by a closed subset of  $C(\Sigma)$  which maps continuously to  $C^{\text{or}}(\text{O})$ . In fact we will choose  $\eta_0$  to be a representative of  $\gamma_0$  which is as simple as possible.

Anyways, with  $A$  as in Proposition 3.3 let

$$Q(A) \subset \mathcal{G}(\Sigma)$$

be the set of unit speed geodesics  $\alpha$  in  $(\tilde{\Sigma}, \rho)$  with the property that the composition of the maps

$$\mathbb{R} \xrightarrow{\alpha} (\tilde{\Sigma}, \rho) \longrightarrow (\hat{\Sigma}, \rho)^c \longrightarrow (\tilde{\text{O}}, \rho_{\text{hyp}})$$

is an  $A$ -quasigeodesic. In these terms, Proposition 3.3 asserts that if  $\alpha : \mathbb{R} \rightarrow (\tilde{\Sigma}, \rho)$  is a unit speed geodesic whose image in  $(\hat{\Sigma}, \rho)$  is as simple as possible, then  $\alpha \in Q(A)$ . Recalling now that by Lemma 3.1 the homotopy class of every closed primitive and essential geodesic  $\gamma_0$  in  $\text{O}$  can be represented by a  $\rho$ -geodesic  $\eta_0 : \mathbb{S}^1 \rightarrow (\Sigma, \rho)$  which is as simple as possible, and that from Lemma 3.2 we get that the property of being as simple as possible is mapping class group invariant, then we get the following fact that we state as a lemma for later reference:

**Lemma 4.1.** *If  $\eta_0 : \mathbb{S}^1 \rightarrow (\Sigma, \rho)$  is any essential  $\rho$ -geodesic which when considered as a map into  $\text{O}$  is as simple as possible, then the measure*

$$\sum_{\eta \in \text{PMap}(\Sigma) \cdot \eta_0} \delta_{\frac{1}{L} \eta}$$

is supported by  $Q(A)$  for all  $L > 0$ .

Now, the fact that the quasigeodesic constant  $A$  is fixed implies that  $Q(A)$  is a closed subset of  $\mathcal{G}(\Sigma)$ . Recall also that every  $A$ -quasigeodesic in  $\tilde{\text{O}}$ , and in particular the image under  $\tilde{\Sigma} \rightarrow \hat{\Sigma} \hookrightarrow \tilde{\text{O}}$  of each element of  $Q(A)$ , is at bounded distance of a  $\rho_{\text{hyp}}$ -geodesic where the bound just depends on  $A$ . In this way we get a continuous map

$$Q(A) \rightarrow \mathcal{G}(\tilde{\text{O}}) \quad (4.2)$$

equivariant under the homomorphism  $\pi_1(\Sigma) \rightarrow \pi_1^{\text{or}}(\text{O}) = \Gamma$ . Now, pushing currents forward with (4.2) (at the end of the day currents are measures) we get a continuous map

$$\Pi : \{\lambda \in C(\Sigma) \text{ supported by } Q(A)\} \rightarrow C^{\text{or}}(\text{O}). \quad (4.3)$$

from the closed subset of  $C(\Sigma)$  consisting of currents supported by the closed set  $Q(A)$  to the space of currents on  $\text{O}$ .

The map  $\Pi$  given in (4.3) induces in turn a continuous map

$$\Pi_* : \text{measures on } \left\{ \begin{array}{l} \lambda \in C(\Sigma) \\ \text{supported by } Q(A) \end{array} \right\} \rightarrow \text{measures on } C^{\text{or}}(\text{O}) \quad (4.4)$$

from the space of Radon measures on  $\{\lambda \in C(\Sigma) \text{ supported by } Q(A)\}$  to the space of Radon measures on  $C^{\text{or}}(\text{O})$ .

For  $\gamma_0$  as in the statement of Theorem 1.2, let  $\eta_0$  be as provided by Lemma 3.1, and note that we get from Lemma 4.1 that the measure  $\sum_{\eta \in \text{Map}(\Sigma) \cdot \eta_0} \delta_{\frac{1}{L}} \eta$  is supported by the domain of (4.3). Its image under (4.4) is in fact nothing other than the measure  $\sum_{\gamma \in \text{Map}^{\text{or}}(\text{O}) \cdot \gamma_0} \delta_{\frac{1}{L}} \gamma$ . Applying  $\Pi_*$  to both sides of the limit in Mirzakhani’s counting theorem we get:

$$\begin{aligned} \Pi_* \left( \frac{c^{\text{PMap}}(\eta_0)}{\mathfrak{b}_{g,r}^{\text{PMap}}} \cdot \mathfrak{m}_{\text{Thu}}^\Sigma \right) &= \Pi_* \left( \lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} \sum_{\eta \in \text{PMap}(\Sigma) \cdot \eta_0} \delta_{\frac{1}{L}} \eta \right) \\ &= \lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} \Pi_* \left( \sum_{\eta \in \text{PMap}(\Sigma) \cdot \eta_0} \delta_{\frac{1}{L}} \eta \right) \\ &= \lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} \sum_{\gamma \in \text{PMap}^{\text{or}}(\text{O}) \cdot \gamma_0} \delta_{\frac{1}{L}} \gamma. \end{aligned} \tag{4.5}$$

Now the  $\text{Map}^{\text{or}}(\text{O})$ -orbit of  $\gamma_0$  is a disjoint union of orbits under  $\text{PMap}^{\text{or}}(\text{O})$ , more precisely of  $\frac{|\text{Map}^{\text{or}}(\text{O})/\text{PMap}^{\text{or}}(\text{O})|}{|\text{Stab}_{\text{Map}^{\text{or}}(\text{O})}(\gamma_0)/\text{Stab}_{\text{PMap}^{\text{or}}(\text{O})}(\gamma_0)|}$  orbits. Theorem 1.2 follows when we apply (4.5) to each one of these orbits and we set

$$C(\gamma_0) = \frac{|\text{Map}^{\text{or}}(\text{O})/\text{PMap}^{\text{or}}(\text{O})|}{|\text{Stab}_{\text{Map}^{\text{or}}(\text{O})}(\gamma_0)/\text{Stab}_{\text{PMap}^{\text{or}}(\text{O})}(\gamma_0)|} \cdot \frac{c^{\text{PMap}}(\eta_0)}{\mathfrak{b}_{g,r}^{\text{PMap}}}$$

and

$$\mathfrak{m}_{\text{Thu}} = \Pi_* \left( \mathfrak{m}_{\text{Thu}}^\Sigma \right).$$

We have proved Theorem 1.2.

Rather, we have proved Theorem 1.2 while assuming Proposition 3.3. Anyways, before proving the proposition let us prove the other theorems mentioned in the introduction and comment briefly on the measure  $\mathfrak{m}_{\text{Thu}}$ .

#### 4.4 | A comment on the measure $\mathfrak{m}_{\text{Thu}}$

In the course of the proof of Theorem 1.2 we identified the measure  $\mathfrak{m}_{\text{Thu}}$  as the push-forward of the Thurston measure  $\mathfrak{m}_{\text{Thu}}^\Sigma$  associated to  $\Sigma$  under the map  $\Pi_*$ . We give now a slightly more intrinsic interpretation of this measure. Note first that every simple essential geodesic in  $\Sigma$  is as simple as possible in  $\text{O}$ . This means that the set  $\mathbb{R}_{\geq 0} \cdot \mathcal{ML}_\Sigma(\Sigma)$  of multiples of integral measured laminations on  $\Sigma$  is contained in  $\mathcal{Q}(A)$ . Since the latter is closed, we also have that the full space of measured laminations on  $\Sigma$  is contained in  $\mathcal{Q}(A)$ , that is,  $\mathcal{ML}(\Sigma) \subset \mathcal{Q}(A)$ . We thus get from the construction (4.1) of the Thurston measure  $\mathfrak{m}_{\text{Thu}}^\Sigma$  that

$$\mathfrak{m}_{\text{Thu}} = \lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} \sum_{\gamma \in \mathcal{ML}_\Sigma(\Sigma)} \delta_{\frac{1}{L}} \hat{\gamma}, \tag{4.6}$$

where, for lack of better notation, we let  $\hat{\gamma}$  be the geodesic representative in  $\text{O}$  of the homotopy class represented by  $\gamma \in \mathcal{ML}_\Sigma(\Sigma)$ . The multicurve curve  $\hat{\gamma}$  is simple in the sense that its lifts to

the universal cover  $\tilde{O}$  never cross each other. If we denote by  $\mathcal{ML}_{\mathbb{Z}}(O)$  the set of, in this sense, simple geodesic multicurves in  $O$ , then we can rewrite (4.6) as

$$\mathbf{m}_{\text{Thu}} = \lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} \sum_{\dot{\gamma} \in \mathcal{ML}_{\mathbb{Z}}(O)} \delta_{\frac{1}{L}} \dot{\gamma}. \quad (4.7)$$

Yet another description of  $\mathbf{m}_{\text{Thu}}$  can be given when we recall that  $O$  has a finite normal cover (in the category of orbifolds) which is a surface. This means that there are a hyperbolic surface  $S$  and a finite group  $H$  acting on  $S$  by isometries such that  $O = S/H$ . The cover  $\pi : S \rightarrow O$  induces a bijection between elements in  $\mathcal{ML}_{\mathbb{Z}}(O)$  and the set  $\mathcal{ML}_{\mathbb{Z}}(S)^H$  of  $H$ -invariant simple multicurves in  $S$ . The reader familiar with the construction of the usual Thurston measure for surfaces will have no difficulty proving that the limit

$$\mathbf{m}_{\text{Thu}}^{S,H} \stackrel{\text{def}}{=} \lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} \sum_{\gamma \in \mathcal{ML}_{\mathbb{Z}}(S)^H} \delta_{\frac{1}{L}} \gamma$$

exists, where  $g$  and  $r$  are still the genus and the sum of the numbers of singular points and boundary components of the orbifold  $O$ . Taking into account the identification between  $\mathcal{ML}_{\mathbb{Z}}(O)$  and  $\mathcal{ML}_{\mathbb{Z}}(S)^H$ , we get

$$\mathbf{m}_{\text{Thu}} = \mathbf{m}_{\text{Thu}}^{S,H}.$$

Here the left measure lives in the space  $\mathcal{G}(\tilde{O})$  of geodesics on the universal cover of  $O$ , and the right one lives in the space  $\mathcal{G}(\tilde{S})$  of geodesics in the universal cover of  $S$ , and where the equality makes sense because  $\tilde{O} = \tilde{S}$ .

*Remark.* It also seems probable that one can recover the measure  $\mathbf{m}_{\text{Thu}}^{S,H}$  as a multiple of the measure on  $\mathcal{ML}(S)^H$  obtained by taking a suitable power of the restriction to that subspace of the Thurston symplectic form on  $\mathcal{ML}(S)$ . It would be interesting to do as in [16] and figure out the precise multiple.

Anyways, the reader having just the present paper in mind can ignore these past comments and just continue thinking of  $\mathbf{m}_{\text{Thu}}$  as given in the proof of Theorem 1.2.

## 4.5 | Actually counting curves

Now we prove Theorem 1.1 and Theorem 1.3 from the introduction. Let us start with the latter.

**Theorem 1.3.** *Let  $O$  be a compact orientable hyperbolic orbifold with possibly empty totally geodesic boundary and let  $C^{\text{or}}(O)$  be the associated space of geodesic currents. Then the limit*

$$\lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} |\{\gamma \text{ of type } \gamma_0 \text{ with } F(\gamma) \leq L\}|$$

*exists and is positive for any  $\gamma_0 \in S^{\text{or}}(O)$  and any positive, homogenous, continuous function  $F : C^{\text{or}}(O) \rightarrow \mathbb{R}_{\geq 0}$ . Here  $g$  is the genus of the orbifold  $O$ , and  $r$  is the sum of the numbers of singular points and boundary components.*

The proof of this theorem is the same as that of the analogous result in the case of surfaces [5, 7, 17] but let us recap the argument anyways. Other than the fact that  $\mathbf{m}_{\text{Thu}}$  is a Radon measure, and as such locally finite, we will also need that

$$\mathbf{m}_{\text{Thu}}(t \cdot U) = t^{6g-6+2r} \mathbf{m}_{\text{Thu}}(U) \tag{4.8}$$

for all  $U \subset C^{\text{or}}(\text{O})$  and all  $t \geq 0$ . This equality holds true because it does so for the standard Thurston measure  $\mathbf{m}_{\text{Thu}}^\Sigma$  and because the map (4.3) is homogeneous:  $\Pi(t \cdot \lambda) = t \cdot \Pi(\lambda)$ . Alternatively (4.8) also follows directly from (4.7). Anyways, we are now ready to prove the theorem:

*Proof.* Noting that there is nothing to prove if  $\text{O}$  is exceptional, suppose that this is not the case.

For any such homogenous function  $F : C^{\text{or}}(\text{O}) \rightarrow \mathbb{R}$  we have

$$\begin{aligned} \frac{|\{\gamma \text{ of type } \gamma_0 \text{ with } F(\gamma) \leq L\}|}{L^{6g-6+2r}} &= \left( \frac{1}{L^{6g-6+2r}} \sum_{\gamma \in \text{Map}^{\text{or}}(\text{O}) \cdot \gamma_0} \delta_\gamma \right) (\{F \leq L\}) \\ &= \left( \frac{1}{L^{6g-6+2r}} \sum_{\gamma \in \text{Map}^{\text{or}}(\text{O}) \cdot \gamma_0} \delta_{\frac{1}{L}\gamma} \right) (\{F \leq 1\}). \end{aligned}$$

Now, by Theorem 1.2 the measures in the last line converge, when  $L \rightarrow \infty$ , to the measure  $C(\gamma_0) \cdot \mathbf{m}_{\text{Thu}}$ . Noting that local finiteness of  $\mathbf{m}_{\text{Thu}}$  together with (4.8) implies that

$$C(\gamma_0) \cdot \mathbf{m}_{\text{Thu}}(\{F = 1\}) = 0,$$

we get that

$$C(\gamma_0) \cdot \mathbf{m}_{\text{Thu}}(\{F \leq 1\}) = \lim_{L \rightarrow \infty} \left( \frac{1}{L^{6g-6+2r}} \sum_{\gamma \in \text{Map}^{\text{or}}(\text{O}) \cdot \gamma_0} \delta_{\frac{1}{L}\gamma} \right) (\{F \leq 1\}).$$

Taking all of this together we obtain that

$$C(\gamma_0) \cdot \mathbf{m}_{\text{Thu}}(\{F \leq 1\}) = \lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} |\{\gamma \text{ of type } \gamma_0 \text{ with } F(\gamma) \leq L\}|,$$

and we are done. □

We come now to Theorem 1.1.

**Theorem 1.1.** *Let  $\Gamma \subset \text{PSL}_2 \mathbb{R}$  be a non-elementary finitely generated discrete subgroup and  $\text{O} = \mathbb{H}^2/\Gamma$  the associated 2-dimensional hyperbolic orbifold. Then the limit*

$$\lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} |\{\gamma \text{ of type } \gamma_0 \text{ with } \ell_{\text{O}}(\gamma) \leq L\}|$$

exists and is positive for any  $\gamma_0 \in S^{\text{or}}(\text{O})$ . Here  $g$  is the genus of the orbifold  $\text{O}$  and  $r$  is the sum of the numbers of singular points and ends.

*Proof.* We might once again assume that  $\text{O}$  is not exceptional. Let then  $\bar{\text{O}}$  be a compact hyperbolic orbifold with interior homeomorphic to  $\text{O}$ , consider  $\gamma_0 \in S^{\text{or}}(\text{O})$  as an element in  $S^{\text{or}}(\bar{\text{O}})$ , and apply Theorem 1.2 to get

$$\lim_{L \rightarrow \infty} \frac{1}{L^{6g-6+2r}} \sum_{\gamma \in \text{Map}^{\text{or}}(\bar{\text{O}}) \cdot \gamma_0} \delta_{\frac{1}{L}\gamma} = C(\gamma_0) \cdot \mathbf{m}_{\text{Thu}}. \tag{4.9}$$

Now note that the same argument that proves it for surfaces shows that there is a compact subset  $K \subset \bar{\text{O}} \setminus \partial\bar{\text{O}}$  which contains the geodesic  $\phi(\gamma_0)$  for all  $\phi \in \text{Map}^{\text{or}}(\bar{\text{O}})$  (see, for example, [5, 7]). It follows that the measures in (4.9) are all supported by the set  $C_K^{\text{or}}(\bar{\text{O}})$  of currents in  $C^{\text{or}}(\text{O})$  whose support projects to a subset of  $K$ . Now, as was first proved by Bonahon [2] (see also [7, Exercise 3.9]) we have that hyperbolic length function  $\ell_{\text{O}}$  extends continuously to  $C_K^{\text{or}}(\bar{\text{O}})$ . The claim of Theorem 1.1 follows when we repeat word-by-word the argument in the proof of Theorem 1.3.  $\square$

## 5 | THE KEY OBSERVATIONS

In this section we get the tools needed to prove Proposition 3.3 in the next section. Notation will be as in the proposition:  $\text{O} = \bar{\text{O}}/\Gamma$  is a compact orientable hyperbolic orbifold with possibly empty totally geodesic boundary,

$$\hat{\Sigma} = \bar{\text{O}} \setminus \mathcal{N}_{\text{hyp}}(\text{sing}(\Gamma), \delta)$$

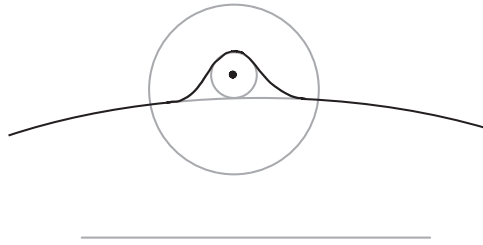
is as in (2.1), and  $\rho$  is the metric on  $\hat{\Sigma}$  constructed in Section 2.7.

### 5.1 | Choosing $\delta$

So far, the only condition we have imposed on  $\delta > 0$  is that it is smaller than  $\frac{1}{3}\epsilon$  where  $\epsilon > 0$  satisfies conditions (C1)–(C3) from Section 2.5. We are momentarily going to give a more stringent condition on  $\delta$ , but first recall that the convex hull of a connected set  $X$  in a negatively curved manifold is the smallest closed connected set  $X'$  with the following property: *any path in  $X$  is homotopic relative to its endpoints to a geodesic path contained in  $X'$* . With this language we fix  $\delta$  so that, with  $\epsilon > 0$  as fixed in Section 2.5, the following holds whenever  $\gamma \subset \bar{\text{O}}$  is a  $\rho_{\text{hyp}}$ -geodesic segment:

$$\begin{aligned} &\text{If } r \geq 50\epsilon \text{ and if } p \in \bar{\text{O}} \text{ is such that } d_{\text{hyp}}(p, \mathcal{N}_{\text{hyp}}(\gamma, r)) \leq 2\delta, \\ &\text{then the } \rho_{\text{hyp}}\text{-convex hull of } \mathcal{N}_{\text{hyp}}(\gamma, r) \cup B_{\text{hyp}}(p, 2\delta) \text{ is contained} \\ &\text{in } \mathcal{N}_{\text{hyp}}(\gamma, r) \cup B_{\text{hyp}}(p, \frac{\epsilon}{2}). \end{aligned} \tag{C4}$$

The lower bound on  $r$  guarantees that  $\mathcal{N}_{\text{hyp}}(\gamma, r)$  is uniformly convex. In particular, the existence of such a  $\delta$  is evident when one considers the limit case  $\delta = 0$ . In any case, a computation shows that any  $\delta < (\frac{\epsilon}{4})^3$  works. See Figure 3 for a schematic representation of (C4).



**FIGURE 3** Schematic representation of (C4). If a point is close enough to the boundary of the uniformly convex set  $\mathcal{N}_{\text{hyp}}(\gamma, r)$ , then its distance to the horizon is really small. The darker line represents the boundary of the convex hull of  $\mathcal{N}_{\text{hyp}}(\gamma, r) \cup B_{\text{hyp}}(p, 2\delta)$ .

The reason for imposing (C4) will be clear shortly, but first we need some more notation. If  $\gamma \subset \tilde{O}$  is a  $\rho_{\text{hyp}}$ -geodesic (always compact) segment, consider for  $r > 0$  the set

$$\mathcal{P}(\gamma, r) = \{p \in \text{sing}(\Gamma) \text{ with } d_{\text{hyp}}(p, \gamma) \leq r + 2\delta\} \tag{5.1}$$

of points which are at most at distance  $2\delta$  from the  $r$ -neighborhood  $\mathcal{N}_{\text{hyp}}(\gamma, r)$  around  $\gamma$ , and for  $t > 0$  let

$$\mathcal{U}(\gamma, r, t) = \mathcal{N}_{\text{hyp}}(\gamma, r) \cup \left( \bigcup_{p \in \mathcal{P}(\gamma, r)} B_p(t, d_{\text{hyp}}) \right) \tag{5.2}$$

be the union of that  $r$ -neighborhood and the  $t$ -balls around each point in  $\mathcal{P}(\gamma, r)$ . The following lemma gives us, for  $t = 2\delta$ , some control of the convex hull of this set with respect to the metric  $\rho$ :

**Lemma 5.1.** *If  $\gamma \subset \tilde{O}$  is a  $\rho_{\text{hyp}}$ -geodesic segment, then we have*

$$(\rho\text{-convex hull of } \mathcal{U}(\gamma, r, 2\delta) \cap \hat{\Sigma}) \subset \left( \mathcal{U}\left(\gamma, r, \frac{\epsilon}{2}\right) \cap \hat{\Sigma} \right)$$

for all  $r \geq 50\epsilon$ . Here  $\mathcal{U}(\gamma, r, t)$  is as in (5.2).

Before launching the proof recall that convexity of a closed set  $X$  is a local property of its boundary. We thus get the following useful property (Figure 4):

If  $(X_0, X_1, \dots)$  is a countable locally finite collection of closed subsets of a negatively curved manifold with  $X_0 \cup X_i$  convex for all  $i$  and with  $X_i \cap X_j = \emptyset$  for all  $i \neq j \geq 1$ , then  $\cup_{i=0}^{\infty} X_i$  is convex. (\*)

We now prove the lemma.

*Proof.* Set  $X_0 = \mathcal{N}_{\text{hyp}}(\gamma, r)$  and for  $p \in \mathcal{P}(\gamma, r)$  let  $Y_p$  be the  $\rho_{\text{hyp}}$ -convex hull of the union of  $\mathcal{N}_{\text{hyp}}(\gamma, r) \cap \hat{\Sigma}$  and  $B_{\text{hyp}}(p, 2\delta)$ . Since  $r \geq 50\epsilon$  then we get from (C4) that  $X_p = Y_p \setminus X_0 \subset B_{\text{hyp}}(p, \frac{\epsilon}{2})$ . This implies that the collection of sets  $\{X_p \text{ with } p \in \mathcal{P}(\gamma, r)\}$  is locally finite and that

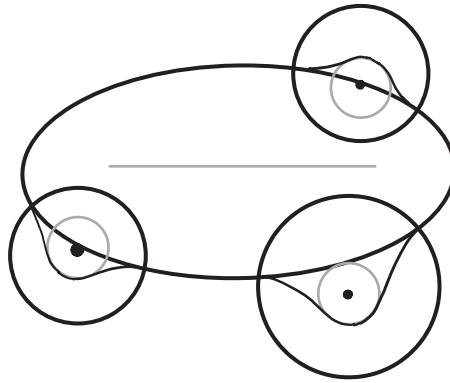


FIGURE 4 Schematic representation of Lemma 5.1

$X_p \cap X_q = \emptyset$  for all distinct  $p, q \in \mathcal{P}(\gamma, r)$ . We get thus from (\*) that

$$X = X_0 \cup \left( \bigcup_{p \in \mathcal{P}(\gamma, r)} X_p \right)$$

is  $\rho_{\text{hyp}}$ -convex, meaning that its boundary is  $\rho_{\text{hyp}}$ -convex. However, we have by construction that

$$\partial X \subset \mathcal{N}_{\text{hyp}}(\gamma, r) \cup \left( \bigcup_{p \in \mathcal{P}(\gamma, r)} \left( B_{\text{hyp}}(p, \frac{\epsilon}{2}) \setminus B_{\text{hyp}}(p, 2\delta) \right) \right),$$

which means that  $\partial X$  is not only contained in  $\hat{\Sigma}$  but even contained in the part of  $\hat{\Sigma}$  where the metrics  $\rho$  and  $\rho_{\text{hyp}}$  agree. This means that  $\partial X$  is not only  $\rho_{\text{hyp}}$ -convex but also  $\rho$ -convex. It follows that  $X \cap \hat{\Sigma}$  is a  $\rho$ -convex set containing  $\mathcal{U}(\gamma, r, 2\delta) \cap \hat{\Sigma}$  but contained in  $\mathcal{U}(\gamma, r, \frac{\epsilon}{2}) \cap \hat{\Sigma}$ . The claim follows.  $\square$

## 5.2 | Heights and outgoing rays

Continuing with the same notation, let  $\gamma \subset \tilde{\mathcal{O}}$  be a  $\rho_{\text{hyp}}$ -geodesic segment and let  $p \in \tilde{\mathcal{O}} \setminus \gamma$  be a point not on  $\gamma$ . Under the  $\gamma$ -outgoing ray at  $p$  we understand the  $\rho_{\text{hyp}}$ -geodesic ray starting at  $p$  in the direction of the gradient of the function  $d_{\text{hyp}}(\gamma, \cdot)$  — that is, the ray that  $p$  would follow to escape from  $\gamma$  at the fastest possible rate.

Now let  $\eta$  be a  $\rho$ -geodesic segment in  $\hat{\Sigma}$  whose endpoints lie on  $\gamma$ . The  $\gamma$ -height of  $\eta$

$$h_\gamma(\eta) = \max \left\{ d_{\text{hyp}}(\gamma, p) \mid \begin{array}{l} \text{where } p \in \text{sing}(\Gamma) \setminus \gamma \text{ is such that} \\ \text{the } \gamma\text{-outgoing ray at } p \text{ meets } \eta \end{array} \right\}$$

is the maximum hyperbolic distance from  $\gamma$  to a cone point  $p \in \text{sing}(\Gamma)$  whose  $\gamma$ -outgoing ray intersects  $\eta$  — here we take  $h_\gamma(\eta) = 0$  if we are taking the maximum over the empty set.

The following lemma asserts that the  $\gamma$ -height of  $\eta$  agrees, up to a small error, with the maximal  $d_{\text{hyp}}$ -distance to  $\gamma$  from points in  $\eta$ .

**Lemma 5.2.** *Let  $\gamma \subset \tilde{O}$  and  $\eta \subset \hat{\Sigma}$  be a  $\rho_{\text{hyp}}$ -geodesic segment and a  $\rho$ -geodesics segment, both with the same endpoints. Then we have*

$$\eta \subset \mathcal{N}_{\text{hyp}}(\gamma, r),$$

where  $r = \max\{50\epsilon, h_\gamma(\eta)\} + \epsilon$ .

*Proof.* Let  $Q$  be the set of those singular points  $p \in \text{sing}(\Gamma)$  whose  $\gamma$ -outgoing ray meets  $\eta$ . Set  $r_0 = \max\{50\epsilon, h_\gamma(\eta)\}$  and note that  $Q \subset \mathcal{P}(\gamma, r_0)$ . The geodesic  $\eta$  is homotopic in  $\hat{\Sigma}$  and while fixing its endpoints to a curve contained in

$$\mathcal{U}(\gamma, r_0, 2\delta) \cap \hat{\Sigma} = \left( \mathcal{N}_{\text{hyp}}(\gamma, r_0) \cup \left( \bigcup_{p \in Q} B_{\text{hyp}}(p, 2\delta) \right) \right) \cap \hat{\Sigma}.$$

The  $\rho$ -geodesic  $\eta$  is then contained in the  $\rho$ -convex hull of  $\mathcal{U}(\gamma, r_0, 2\delta) \cap \hat{\Sigma}$  and hence in  $\mathcal{U}(\gamma, r_0, \frac{\epsilon}{2}) \cap \hat{\Sigma} \subset \mathcal{N}_{\text{hyp}}(\gamma, r_0 + \epsilon)$  by Lemma 5.1. We are done. □

### 5.3 | The main observation

Our next goal is to establish the following fact.

**Lemma 5.3.** *Let  $\gamma \subset \tilde{O}$  and  $\eta \subset \hat{\Sigma}$  be, respectively, a  $\rho_{\text{hyp}}$ -geodesic segment and a simple  $\rho$ -geodesic segment, such that both segments have the same endpoints  $\partial\gamma = \partial\eta$ . Suppose that at least one of the following holds:*

- (a)  $h_\gamma(\eta) > 1$ , or
- (b)  $\ell_{\text{hyp}}(\gamma) < \epsilon$  and  $h_\gamma(\eta) > 50\epsilon$ .

*Then there is  $g \in \Gamma \setminus \text{Id}$  such that  $\eta$  and  $g\eta$  transversely intersect at least twice.*

*Remark.* It follows by a limiting argument and Lemma 5.3 that, if we replace the “max” in the definition of  $h_\gamma(\eta)$  by a “sup,” then the lemma also holds when  $\gamma$  and  $\eta$  are a complete  $\rho_{\text{hyp}}$ -geodesic and a complete simple  $\rho$ -geodesic which have the same endpoints in  $\partial_\infty \tilde{O}$ , the boundary at infinity of  $\tilde{O}$ . To see that this is the case parametrize  $\eta : \mathbb{R} \rightarrow \hat{\Sigma}$ , let  $\gamma_n$  be the  $\rho_{\text{hyp}}$ -geodesic segment with endpoints  $\eta(-n)$  and  $\eta(n)$ , set  $\eta_n = \eta([-n, n])$ , and note that

$$h_\gamma(\eta) \leq \liminf_{n \rightarrow \infty} h_{\gamma_n}(\eta_n).$$

It thus follows from the lemma that if  $h_\gamma(\eta) > 1$ , then there are  $n > 0$  and  $g \in \Gamma \neq \text{Id}$  such that  $\eta_n$  and  $g\eta_n$  transversely intersect at least twice. This means a fortiori that  $\eta$  and  $g\eta$  also meet transversely at least twice.

*Proof.* Starting with the proof of Lemma 5.3, suppose that  $\gamma$  and  $\eta$  satisfy one of the two possible conditions in the statement. As a first observation note that if  $\gamma$  and  $\eta$  satisfy (a) (respectively, (b)) and meet in a point other than in their end points, then there are subsegments  $\gamma' \subset \gamma$  and  $\eta' \subset \eta$  with  $\partial\gamma' = \partial\eta'$ , which still satisfy (a) (respectively, (b)) and such that  $\gamma'$  and  $\eta'$  meet only at their



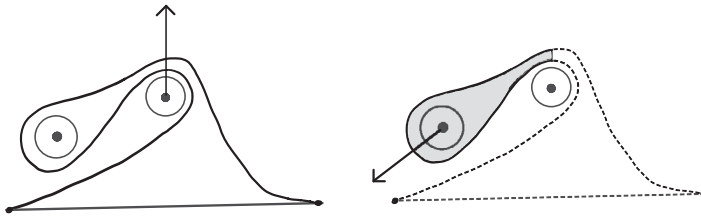


FIGURE 5 Proof of Claim 1

endpoints. This means that we can assume without loss of generality that the loop obtained by concatenating  $\gamma$  and  $\eta$  is simple. Or said differently that  $\gamma$  and  $\eta$  bound a disk  $\Delta$  in  $\mathbb{H}^2$ .

*Claim 1.* There are a  $\rho_{\text{hyp}}$ -geodesic segment  $\bar{\gamma}$  and a subsegment  $\bar{\eta} \subset \eta$  satisfying the following properties.

- (1) The two segments  $\bar{\gamma}$  and  $\bar{\eta}$  have the same endpoints and disjoint interiors.
- (2) The pair  $(\bar{\gamma}, \bar{\eta})$  satisfies one of the two conditions (a) and (b) in the statement of the lemma.
- (3) The disk  $\bar{\Delta}$  with boundary  $\bar{\gamma} \cup \bar{\eta}$  contains a point  $p \in \text{sing}(\Gamma) \cap \bar{\Delta}$  with  $d_{\text{hyp}}(p, \bar{\gamma}) = h_{\bar{\gamma}}(\bar{\eta})$ .

We suggest that at first, instead of studying the proof of the claim, the reader spends some time looking at Figure 5.

*Proof of Claim 1.* If the disk  $\Delta$  bounded by the concatenation of  $\gamma$  and  $\eta$  satisfies (3), then we have nothing to prove. If this is not the case, then we will find a hyperbolic geodesic segment  $\gamma'$  and a subsegment  $\eta' \subset \eta$  satisfying (1) and (2), and such that  $\eta'$  is at least  $\epsilon$ -shorter than  $\eta$ , that is,  $\ell_{\rho}(\eta') \leq \ell_{\rho}(\eta) - \epsilon$ . Now, if the disk  $\Delta'$  associated to  $\gamma'$  and  $\eta'$  satisfies (3), then we are done. Otherwise we iterate our procedure. But this process can only be repeated finitely many times because at each step we lose a definite amount of length, and the length of the original segment  $\eta$  is finite.

Let us see how we find  $\gamma'$  and  $\eta'$ . We start by taking a point  $p \in \text{sing}(\Gamma)$  such that the  $\gamma$ -outgoing ray  $\sigma$  at  $p$  meets  $\eta$  and with  $d_{\text{hyp}}(p, \gamma) = h_{\gamma}(\eta)$ . Note that Lemma 5.2 implies that all intersections of  $\sigma$  and  $\eta$  happen in the annulus  $B_{\text{hyp}}(p, \epsilon) \cap \hat{\Sigma} = B_{\text{hyp}}(p, \epsilon) \setminus B_{\text{hyp}}(p, \delta)$ . Since the outgoing ray  $\sigma$  intersects  $\eta$  and we are assuming that  $p \notin \Delta$ ,  $\sigma$  must intersect  $\eta$  at least twice. We deduce that there is a closed subsegment

$$\gamma' \subset \sigma \cap (B_{\text{hyp}}(p, \epsilon) \cap \hat{\Sigma})$$

with  $\gamma' \cap \eta = \partial\gamma'$ . Let  $\eta' \subset \eta$  be the subsegment of  $\eta$  bounded by  $\partial\gamma' \subset \eta$ .

By construction the pair  $\gamma', \eta'$  satisfies (1). Moreover, since  $\eta$  has to travel at least distance  $49\epsilon$  to go from  $\gamma$  to  $B_{\text{hyp}}(p, \epsilon)$  and since the metrics  $\rho$  and  $\rho_{\text{hyp}}$  agree on  $B_{\text{hyp}}(p, 50\epsilon) \setminus B_{\text{hyp}}(p, \delta)$ , we get that

$$\ell_{\rho}(\eta') \leq \ell_{\rho}(\eta) - 98\epsilon,$$

which beats our established goal of reducing the length by  $\epsilon$  by a proud  $97\epsilon$ . It just suffices to prove that the pair  $(\gamma', \eta')$  satisfies (2), meaning that one of the conditions (a) or (b) holds. Actually, we

are going to argue that they satisfy (b). First,  $\gamma'$  is shorter than  $\epsilon$  by construction. It thus suffices to check that  $h_{\gamma'}(\eta') > 50\epsilon$ . By Lemma 5.2 it suffices to prove that  $\eta'$  exits  $\mathcal{N}_{\text{hyp}}(\gamma', 51\epsilon)$ , or even better, that it exits the ball  $B_{\text{hyp}}(p, 52\epsilon)$ . Indeed, since  $\eta'$  has both endpoints in  $\sigma$ , a hyperbolic ray emanating from  $p$ , since  $\eta'$  is contained in  $\eta$ , a simple geodesic which exists  $B_{\text{hyp}}(p, 52\epsilon)$  in both directions, and since the metrics  $\rho$  and  $\rho_{\text{hyp}}$  agree by construction on  $B_{\text{hyp}}(p, 52\epsilon) \setminus B_p(2\delta, d_{\text{hyp}})$ , we get from Lemma 2.1 that  $\eta'$  cannot be contained in  $B_{\text{hyp}}(p, 52\epsilon)$ . We have proved the claim.  $\square$

Continuing with the proof of Lemma 5.3 and with notation as in Claim 1, choose  $g \in \text{Stab}_{\Gamma}(p)$  with rotation angle  $\theta \in [\frac{2\pi}{3}, \frac{4\pi}{3}]$ . Note that such  $g$  always exists.

*Claim 2.* We have  $g^{\pm 1}(\bar{\gamma}) \cap \bar{\Delta} = \emptyset$ .

*Proof of Claim 2.* Suppose first that the pair  $(\bar{\gamma}, \bar{\eta})$  satisfies (a), meaning that

$$d_{\text{hyp}}(p, \bar{\gamma}) = h_{\bar{\gamma}}(\bar{\eta}) > 1.$$

A computation using standard hyperbolic trigonometry implies that

$$d_{\text{hyp}}(\bar{\gamma}, g^{\pm 1}\bar{\gamma}) = 2 \cosh^{-1} \left( \frac{\sqrt{3}}{2} \cosh(d_{\text{hyp}}(p, \bar{\gamma})) \right) \geq \frac{3}{2} d_{\text{hyp}}(p, \bar{\gamma}) > h_{\bar{\gamma}}(\bar{\eta}) + \epsilon.$$

The claim follows thus because  $\bar{\Delta} \subset \mathcal{N}_{\text{hyp}}(\bar{\gamma}, h_{\bar{\gamma}}(\bar{\eta}) + \epsilon)$  by Lemma 5.2. We are done if the pair  $(\bar{\gamma}, \bar{\eta})$  satisfies (a). Suppose now that it satisfies (b) and let  $p^*$  be projection of  $p$  to  $\bar{\gamma}$ . Now, either again a hyperbolic geometry computation, or just plainly comparing with the comparison euclidean triangle one gets that

$$d_{\text{hyp}}(p^*, g^{\pm 1}p^*) \geq \sqrt{3} \cdot d_{\text{hyp}}(p, p^*) \geq \sqrt{3} \cdot h_{\bar{\gamma}}(\bar{\eta}) > h_{\bar{\gamma}}(\bar{\eta}) + 3\epsilon.$$

Again the claim follows from Lemma 5.2.  $\square$

We are now ready to conclude the proof of the lemma. First note that

$$g\bar{\Delta} \cap \bar{\Delta} \neq \emptyset,$$

because  $gp = p$ . Since  $g\bar{\Delta}$  can neither be contained in  $\bar{\Delta}$  nor contain  $\bar{\Delta}$ , we deduce that  $|g(\partial\bar{\Delta}) \cap \partial\bar{\Delta}| \geq 2$ . Since  $\partial\bar{\Delta} = \bar{\gamma} \cup \bar{\eta}$ , we get from Claim 2 that

$$g\bar{\gamma} \cap (\partial\bar{\Delta}) = \emptyset \text{ and } \bar{\gamma} \cap g(\partial\bar{\Delta}) = \emptyset.$$

It follows that  $g(\partial\bar{\Delta}) \cap (\partial\bar{\Delta}) = g\bar{\eta} \cap \bar{\eta}$  and hence that  $|g\bar{\eta} \cap \bar{\eta}| \geq 2$ . The lemma then follows because  $\bar{\eta} \subset \eta$ .  $\square$

### 6 | PROOF OF PROPOSITION 3.3

We are now finally ready to prove the remaining proposition.

**Proposition 3.3.** *Let  $O$  be as in the statement of Theorem 1.2,  $\hat{\Sigma}$  as in (2.1), and  $\rho$  the metric on  $\hat{\Sigma}$  constructed in Section 2.7. There exists  $A \geq 1$  such that any unit speed  $\rho$ -geodesic  $\alpha : \mathbb{R} \rightarrow \hat{\Sigma}$  which is (1) a quasigeodesic in  $(\tilde{O}, \rho_{\text{hyp}})$  and (2) as simple as possible is actually  $A$ -quasigeodesic in  $(\tilde{O}, \rho_{\text{hyp}})$ .*

*Proof.* Note first that compactness of  $\hat{\Sigma}/\Gamma$ , together with  $\Gamma$ -invariance of the metric  $\rho$ , implies that the inclusion  $(\hat{\Sigma}, \rho) \hookrightarrow (\tilde{O}, \rho_{\text{hyp}})$  is locally  $A_0$ -bi-Lipschitz for some  $A_0$ . It follows in particular that for all  $s, t \in \mathbb{R}$  we have

$$A_0 \cdot |s - t| \geq d_{\text{hyp}}(\alpha(s), \alpha(t)). \quad (6.1)$$

The remaining of the proof is devoted to showing that the other inequality in the definition of quasigeodesic holds for some constant independent of the concrete  $\alpha$ .

As a first step we choose a  $\Gamma$ -invariant triangulation  $\mathcal{T}$  of  $\tilde{O}$  whose edges  $\mathcal{T}$  are  $\rho_{\text{hyp}}$ -geodesic segments of length at most  $\epsilon$ . Consider the set

$$\mathcal{P} = \{\text{simplices of } \mathcal{T} \text{ contained in } \mathcal{N}_{\text{hyp}}(\text{sing}(\Gamma), 60\epsilon)\}$$

and let  $|\mathcal{P}| = \cup_{\sigma \in \mathcal{P}} \sigma \subset \tilde{O}$  be the union of all those simplexes. Let also

$$\mathcal{E} = \{\text{edges of } \mathcal{T} \text{ contained in the closure of } \tilde{O} \setminus |\mathcal{P}|\},$$

and let  $|\mathcal{E}| = \cup_{e \in \mathcal{E}} e \subset \tilde{O}$  be the graph obtained by taking the union of all the edges of  $\mathcal{T}$  which are disjoint of the interior of  $|\mathcal{P}|$ . Note that the elements in  $\mathcal{E}$  are not only  $\rho_{\text{hyp}}$ -geodesic but also  $\rho$ -geodesic because the metrics  $\rho$  and  $\rho_{\text{hyp}}$  agree away from a very small neighborhood of  $\text{sing}(\Gamma)$ . Note also that compactness of  $O = \tilde{O}/\Gamma$  and  $\Gamma$ -invariance of  $\mathcal{T}$  imply that there is  $C$  such that the following holds for all  $p \in \tilde{O}$ .

(a) Less than  $C$  elements of  $\mathcal{E}$  intersect the ball  $B_{\text{hyp}}(p, 4)$ .

We care about all of this because we will estimate the  $\rho$ -length of subsegments of  $\alpha$  as in the statement of Proposition 3.3 in terms of the number of edges in  $\mathcal{E}$  that they meet. The key observation is that, since  $\alpha$  as in the statement is as simple as possible and hence also simple, and since it cannot spend infinite time in any single ball since it is quasigeodesic, we get from Lemma 2.1 that there is some  $D > \epsilon$  such that:

(b) if  $\alpha : \mathbb{R} \rightarrow \tilde{O}$  is as in the statement of the proposition and if  $[s, t] \subset \mathbb{R}$  is such that  $\alpha[s, t] \cap |\mathcal{E}| = \emptyset$ , then  $|s - t| < D$ .

We get from (b) that a geodesic  $\alpha$  as in the statement never spends much time without meeting one of the edges in  $\mathcal{E}$ . We prove next that once such an  $\alpha : \mathbb{R} \rightarrow \hat{\Sigma}$  leaves  $e \in \mathcal{E}$ , it never comes back to  $e$ . Indeed, suppose that  $s < t$  are such that  $\alpha(s), \alpha(t) \in e$  for some  $e \in \mathcal{E}$ . Denote by  $\gamma$  the subsegment of  $e$  between  $\alpha(s)$  and  $\alpha(t)$  and let  $\eta = \alpha[s, t]$ . We claim first that  $\eta \subset \mathcal{N}_{\text{hyp}}(\gamma, 55\epsilon)$ . Otherwise we get from Lemma 5.2 that  $h_\gamma(\eta) > 50\epsilon$  and then from Lemma 5.3 that there is  $g \in \Gamma$  such that  $|\eta \cap g\eta| \geq 2$ , contradicting the assumption that  $\alpha$  is as simple as possible. We have thus proved that  $\eta \subset \mathcal{N}_{\text{hyp}}(\gamma, 55\epsilon)$ . But noting that  $\mathcal{N}_{\text{hyp}}(\gamma, 55\epsilon) \subset \mathcal{N}_{\text{hyp}}(e, 55\epsilon) \subset \hat{\Sigma}$  is contractible and that both metrics  $\rho$  and  $\rho_{\text{hyp}}$  agree thereon, we deduce that  $\gamma = \eta$  because both are geodesic segments with the same endpoints. We have established the following key fact.

(c) If  $\alpha : \mathbb{R} \rightarrow \tilde{O}$  is as in the statement of the proposition and if  $s < t \in \mathbb{R}$  are such that  $\alpha(s), \alpha(t) \in e$  for some  $e \in \mathcal{E}$ , then  $\alpha[s, t] \subset e$ .

The reader surely can at this point imagine how (b) and (c) interplay, but might still be wondering why we bothered to state (a) at all. Well, the reason is coming. Still assuming that  $\alpha : \mathbb{R} \rightarrow \hat{\Sigma}$  is a  $\rho$ -geodesic as in the statement, suppose that  $s, t \in \mathbb{R}$  are such that  $d_{\text{hyp}}(\alpha(s), \alpha(t)) \leq 4$ , let  $\gamma \subset \tilde{O}$  be the  $\rho_{\text{hyp}}$ -geodesic segment joining  $\alpha(s)$  and  $\alpha(t)$ , and finally let  $p$  be the midpoint of  $\gamma$ . Since  $\alpha$  is as simple as possible, we get from Lemma 5.3 that  $h_\gamma(\alpha[s, t]) \leq 1$ . Lemma 5.2 yields then that

$$\alpha([s, t]) \subset \mathcal{N}_{\text{hyp}}(\gamma, 1 + \epsilon) \subset B_{\text{hyp}}(p, 4).$$

We thus get from (a) that  $\alpha[s, t]$  meets at most  $C$  elements of  $\mathcal{E}$ , and from (c) that, if we cut  $\alpha[s, t]$  at all the points where we enter an element of  $\mathcal{E}$ , then we produce at most  $C + 1$  segments. Since all of them have at most length  $D$  by (b), we obtain:

(d) If  $\alpha : \mathbb{R} \rightarrow \tilde{O}$  is as in the statement of the proposition and if  $s, t \in \mathbb{R}$  are such that  $d_{\text{hyp}}(\alpha(s), \alpha(t)) \leq 4$ , then  $|s - t| \leq (C + 1) \cdot D$ .

We are almost at the end of the proof of the proposition. Recalling that  $\alpha$  as in the statement of the proposition is a quasigeodesic, let  $\gamma \subset \tilde{O}$  be the hyperbolic geodesic with the same endpoints as  $\alpha$ , and let

$$\pi : \mathbb{H}^2 \rightarrow \gamma$$

be the nearest point projection. From Lemma 5.3, or rather from the comment following the said lemma, we get that  $h_\gamma(\alpha(\mathbb{R})) \leq 1$ . Lemma 5.2 implies that  $d_{\text{hyp}}(\alpha(s), \pi(\alpha(s))) \leq 1 + \epsilon$  for all  $s$ . Now, if we have  $s, t \in \mathbb{R}$  with  $d_{\text{hyp}}(\pi(\alpha(s)), \pi(\alpha(t))) = 1$ , we get  $d_{\text{hyp}}(\alpha(s), \alpha(t)) \leq 3 + 2\epsilon < 4$ . We thus get from (d) that:

(e) if  $\alpha : \mathbb{R} \rightarrow \tilde{O}$  is as in the statement of the proposition, if  $\gamma$  is the  $\rho_{\text{hyp}}$ -geodesic at bounded distance of  $\alpha$ , and if  $\pi : \tilde{O} \rightarrow \gamma$  is the nearest point projection, then we have

$$|s - t| \leq (C + 1) \cdot D \stackrel{\text{def}}{=} A_1$$

for all  $s, t \in \mathbb{R}$  with  $d_{\text{hyp}}(\pi(\alpha(s)), \pi(\alpha(t))) \leq 1$ .

Now, if we have  $s < t \in \mathbb{R}$  arbitrary let  $s_0 = s$  and, as long as  $s_k < t$ , define iteratively  $s_{k+1} > s_k$  as follows:

- If  $d_{\text{hyp}}(\pi(\alpha(s_k)), \pi(\alpha(t))) < 1$ , then set  $s_{k+1} = t$ .
- Else  $s_{k+1} = \max\{s' \in [s, t] \text{ with } d_{\text{hyp}}(\pi(\alpha(s_k)), \pi(\alpha(s'))) = 1\}$ .

In this way we get a sequence

$$s = s_0 < s_1 < s_2 < \dots < s_{N-1} < s_N = t,$$

where  $N$  satisfies

$$d_{\text{hyp}}(\pi(\alpha(s)), \pi(\alpha(t))) \geq N - 1.$$

On the other hand, we get from (e) that  $|s_k - s_{k+1}| \leq A_1$ . This means that

$$N \geq \frac{|s - t|}{A_1}.$$

Taking all of this together we have that for any  $\alpha : \mathbb{R} \rightarrow \hat{\Sigma}$  as in the statement of the Proposition, we will have

$$d_{\text{hyp}}(\alpha(s), \alpha(t)) \geq d_{\text{hyp}}(\pi(\alpha(s)), \pi(\alpha(t))) \geq \frac{|s - t|}{A_1} - 1 \quad (6.2)$$

for all  $s, t \in \mathbb{R}$ .

The claim follows then from (6.1) and (6.2) with  $A = \max\{A_0, A_1, 1\}$ .  $\square$

## ACKNOWLEDGEMENTS

This has been one of those projects that for whatever reason take a long time to be completed. So long in fact that it is impossible to make a comprehensive list of everyone we owe our gratitude to, and wishing not to be unfair we thank nobody — ingen nämnd ingen glömd. With one exception, because the first author has not forgotten that during the start of the project she was supported by Pekka Pankka's [Academy of Finland](#) project #297258 at the University of Helsinki. The first author gratefully acknowledges support from [EPSRC](#) grant EP/T015926/1.

## JOURNAL INFORMATION

The *Transactions of the London Mathematical Society* is wholly owned and managed by the London Mathematical Society, a not-for-profit Charity registered with the UK Charity Commission. All surplus income from its publishing programme is used to support mathematicians and mathematics research in the form of research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

## REFERENCES

1. J. Aramayona and C. J. Leininger, *Hyperbolic structures on surfaces and geodesic currents*, Algorithmic and geometric topics around free groups and automorphisms, Advance Courses in Mathematics, CRM Barcelona, Birkhäuser/Springer, Cham, 2017, pp. 111–149.
2. F. Bonahon, *Bouts des variétés hyperboliques de dimension 3*, Ann. of Math. (2) **124** (1986), no. 1, 71–158.
3. F. Bonahon, *The geometry of Teichmüller space via geodesic currents*, Invent. Math. **92** (1988), no. 1, 139–162.
4. F. Bonahon, *Geodesic currents on negatively curved groups*, Arboreal group theory (Berkeley, CA, 1988), Mathematical Sciences Research Institute Publications, vol. 19, Springer, New York, 1991, pp. 143–168.
5. V. Erlandsson, H. Parlier, and J. Souto, *Counting curves, and the stable length of currents*, J. Eur. Math. Soc. (JEMS) **22** (2020), no. 6, 1675–1702.
6. V. Erlandsson and J. Souto, *Counting curves in hyperbolic surfaces*, Geom. Funct. Anal. **26** (2016), no. 3, 729–777.
7. V. Erlandsson and J. Souto, *Mirzakhani's curve counting and geodesic currents*, Progress in Mathematics, vol. 345, Birkhäuser, Cham, Switzerland, 2022.
8. B. Farb and D. Margalit, *A primer on mapping class groups*, Princeton Mathematical Series, vol. 49, Princeton University Press, Princeton, NJ, 2012.
9. M. Gendulphé, *What's wrong with the growth of simple closed geodesics on nonorientable hyperbolic surfaces*, <https://arxiv.org/abs/1706.08798>, 2017.
10. J. Hass and P. Scott, *Shortening curves on surfaces*, Topology **33** (1994), no. 1, 25–43.
11. M. Magee, *Counting one-sided simple closed geodesics on Fuchsian thrice punctured projective planes*, Int. Math. Res. Not. IMRN **13** (2020), no. 6, Art. ID rnm112.

12. D. Martínez-Granado and D. Thurston, *From curves to currents*, Forum Math. Sigma **9** (2021).
13. M. Mirzakhani, *Simple geodesics on hyperbolic surfaces and the volume of the moduli space of curves*, PhD thesis, ProQuest LLC, Ann Arbor, MI, 2004.
14. M. Mirzakhani, *Growth of the number of simple closed geodesics on hyperbolic surfaces*, Ann. of Math. (2) **168** (2008), no. 1, 97–125.
15. M. Mirzakhani, *Counting mapping class group orbits on hyperbolic surfaces*, Preprint, <https://arxiv.org/abs/1601.03342>, 2016.
16. L. Monin and V. Telpukhovskiy, *On normalizations of Thurston measure on the space of measured laminations*, Topology Appl. **267** (2019), 106878.
17. K. Rafi and J. Souto, *Geodesic currents and counting problems*, Geom. Funct. Anal. **29** (2019), no. 3, 871–889.
18. P. Scott, *The geometries of 3-manifolds*, Bull. London Math. Soc. **15** (1983), no. 5, 401–487.