

QUANTIFYING THE SPARSENESS OF SIMPLE GEODESICS ON HYPERBOLIC SURFACES

Peter Buser, Hugo Parlier*

Abstract. The goal of the article is to provide different explicit quantifications of the non density of simple closed geodesics on hyperbolic surfaces. In particular, we show that within any embedded metric disk on a surface, lies a disk of radius only depending on the topology of the surface (and the size of the first embedded disk), which is disjoint from any simple closed geodesic.

1. INTRODUCTION

The set of simple closed geodesics on finitely generated hyperbolic surfaces has many remarkable properties and is related to various aspects of geometric and dynamical properties of moduli spaces and mapping class groups. Among these properties is a result of Birman and Series which states that the larger set of simple complete geodesics is nowhere dense and has Hausdorff dimension 1, the same result holding true for geodesics with uniformly bounded self-intersection number [2]. This result, which might seem surprising at first in light of other phenomena, answered a question raised by Jorgensen [8] who first exhibited surfaces with non-dense sets of simple closed geodesics.

It is easy to see that this is really a feature of negative curvature: for instance simple complete geodesics on a flat torus leaving from a given point cover the entire surface and the closed ones are dense (even though they form a measure 0 subset of the surface). Of course on a flat torus all closed geodesics are simple and similarly, on a hyperbolic surface, the set of *all* closed geodesics is not only dense but also dense in the unit tangent bundle. This is one of many instances of how simple geodesics are rare within the set of all closed geodesics.

*Research partially supported by ANR/FNR project SoS, INTER/ANR/16/11554412/SoS, ANR-17-CE40-0033, by Swiss National Science Foundation grant number PP00P2_153024 and by U.S. National Science Foundation grants DMS 1107452, 1107263, 1107367 RNMS: Geometric structures And Representation varieties (the GEAR Network).

2010 Mathematics Subject Classification: Primary: 32G15, 30F10. Secondary: 30F60, 53C22, 58J50.

Key words and phrases: hyperbolic surfaces, simple closed geodesics

Another prime example is for the growth of the number of curves on hyperbolic surfaces. By Huber’s asymptotic law [7], the number of closed geodesics of length less than L grows asymptotically like e^L/L . In contrast, Mirzakhani showed that simple geodesics have very different asymptotic growth as they grow polynomially in L with leading term on the order of L^{6g-6+n} where g is the genus and n the number of cusps [9]. Her results show much more than just asymptotic growth and relate the growth function to the underlying moduli space in many ways. In particular the exact asymptotic behavior doesn’t only depend on the topology but also on the underlying geometry of the surface. That the number of simple curves is bounded above by a polynomial function of length was already one of the key arguments in the results of Birman and Series, and the correct rough asymptotic growth was first proved by Rivin [11]. More generally, a polynomial upper bound holds for curves with bounded intersection number, and there has been a flurry of results showing asymptotic growth for such curves in more general contexts [5, 6, 10, 12, 13].

In this article we quantify the non-density of simple complete geodesics. The Birman-Series result tells us that given any hyperbolic surfaces, in any neighborhood of a point, there is a small disk entirely untouched by any simple complete geodesic. The type of questions we aim to answer are: How large can you take that disk to be? What is the size of the largest disk disjoint from all simple complete geodesics?

The set of points of S that lie on a simple complete geodesic will be denoted $\mathcal{BS}(S)$ or simply \mathcal{BS} and we shall sometimes refer to this set as the Birman-Series set. Given a surface S , we can look at the radius of the largest disk on the complement of \mathcal{BS} . Given a moduli space \mathcal{M} , the size of this “maximal gap” is a function over \mathcal{M} and in fact is continuous. Using this and a computation of what happens towards the boundary of moduli spaces, one can show that there is a positive lower bound to G_S which only depends on the topology of the underlying surface (some of the details can be found in [4]). In particular, this shows the existence of a constant $K_g > 0$ such that any closed hyperbolic surface of genus g has a gap of size K_g (or similarly the existence of a constant $K_{g,n} > 0$ such that any genus g surface with n cusps has a gap of size $K_{g,n}$). One might hope to find a universal lower bound on the size of gaps but in fact this is impossible. Take any ε -dense but finite set of closed geodesics on a closed surface. Then a theorem of Scott [15] asserts that there is a finite cover where all lifts of the closed geodesics are simple. In the cover, the simple geodesics reproduce the ε -density.

So as there are no universal positive lower bounds on K_g , one of our underlying goals is to quantify the constant K_g in terms of g . Our first approach to this leads to a precise computation for surfaces in the thin part of moduli space following the natural thick-thin

decomposition of moduli space with the *systole* function. The *systole* $\text{sys}(S)$ of a finite type hyperbolic surface S is the length of a non-trivial curve of minimal length (by non-trivial we mean non-homotopically trivial and non-peripheral to boundary). Surfaces with systole at least $\varepsilon > 0$ are said to lie in the ε -thick part of moduli space. A first step to proving the above theorem is to show that (closed) surfaces with systole below a certain threshold have a gap of that same size.

Theorem 1.1. *Let $a_g = \frac{1}{4 \cdot (4\pi(g-1))^2}$. If $\text{sys}(S) \leq a_g$, then $C_S \geq a_g$.*

The next step is to deal with surfaces with systole length bounded below. To do so we show the following local result which essentially says that given an embedded disk on the surface, there is a quantifiable gap of a certain radius within that disk. The radius only depends on the topology of the underlying surface (and the size of the initial disk of course).

Theorem 1.2. *Assume $s = \min\{\frac{1}{2}\text{sys}(S), \frac{1}{3}\}$. Then for any $\rho \leq s$ and any disk B_ρ of radius ρ in S there exists a point $p \in B_\rho$ such that*

$$\text{dist}(p, \mathcal{BS}) \geq \rho^2 e^{-M(g-1)},$$

where M is an explicit constant that depends only on s .

The constant M can be taken to be $\frac{194}{s^2} \log(\frac{134}{s})$ and together with the result for thin surfaces, this provides a quantifiable lower bound on K_g . Using the same techniques it also holds for (sufficiently thick) surfaces with cusps.

Our final goal is to establish a quantified local result which does not depend on systole length. It implies an explicit bound on the constant $K_{g,n}$ discussed previously.

Theorem 1.3. *Let S be a hyperbolic surface of genus g with n cusps and B_ρ a disk of radius ρ in the ε -thick part of S , where $0 < \rho < \varepsilon \leq 1/3$. Then there exists a point $p \in B_\rho$ such that*

$$\text{dist}(p, \mathcal{BS}) \geq e^{-3^\kappa R},$$

where $\kappa = 3g - 3 + n$ and $R = 2 \log \frac{1}{\rho} + M(g - 1 + n/2)$ with $M = \frac{195}{\varepsilon^2} \log \frac{134}{\varepsilon}$.

Note than in the above, as B_ρ always lies in the ρ -thick part, so in particular the result holds with $\varepsilon = \rho$. Our local estimates without prior assumptions on systole length are weaker. This comes from our method in which we use a sort of classification of short curves: if they are sufficiently short with respect to the other short curves we treat them like cusps and if not we treat them as “short but not too short curves”.

We end this introduction with some images of the Birman-Series set for genus 2 surfaces illustrating its intricacy; although these geodesics are nowhere dense, the gaps are already quite small.

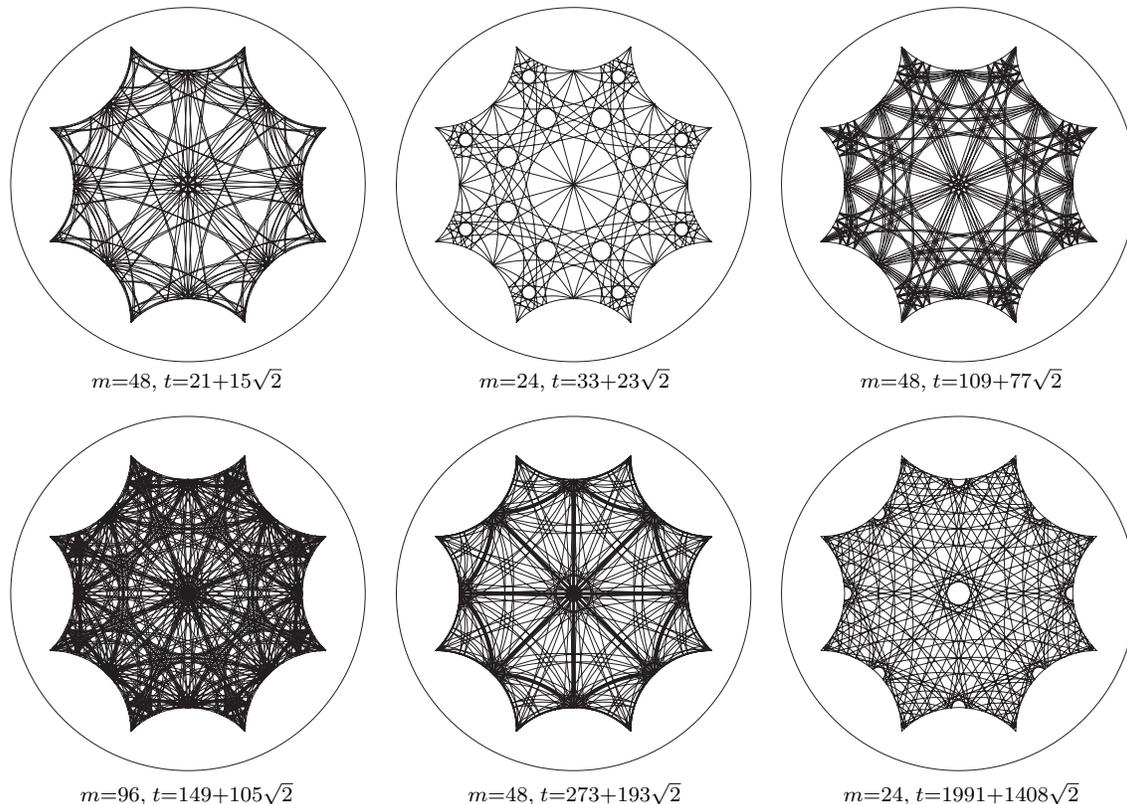


Figure 1: Simple closed geodesics on the Bolza surface

Figure 1 shows a number of simple closed geodesics on the regular fundamental domain of the Bolza surface, the genus 2 Riemann surface with the maximal number of symmetries. Originally the idea was to show all the simple closed geodesics up to length roughly 15 on the same fundamental domain. However, even under optimal printing conditions, the fundamental domain came out evenly black. We have therefore split up the geodesics into families where all members of a family have the same length. Figure 1 shows a few. In this figure, m is the multiplicity, i.e. the number of geodesics in the family, and $t = \operatorname{arccosh}(\ell(\gamma)/2)$ is half the trace of the conjugacy class in the Fuchsian group that represents a closed geodesic of length $\ell(\gamma)$.

Figure 2 shows roughly the first (ordered by length) three hundred simple closed geodesics on two other genus 2 surfaces. These surfaces were obtained by perturbing Fenchel-Nielsen length and twist parameters of the Bolza surface. The size of the largest gaps appears to

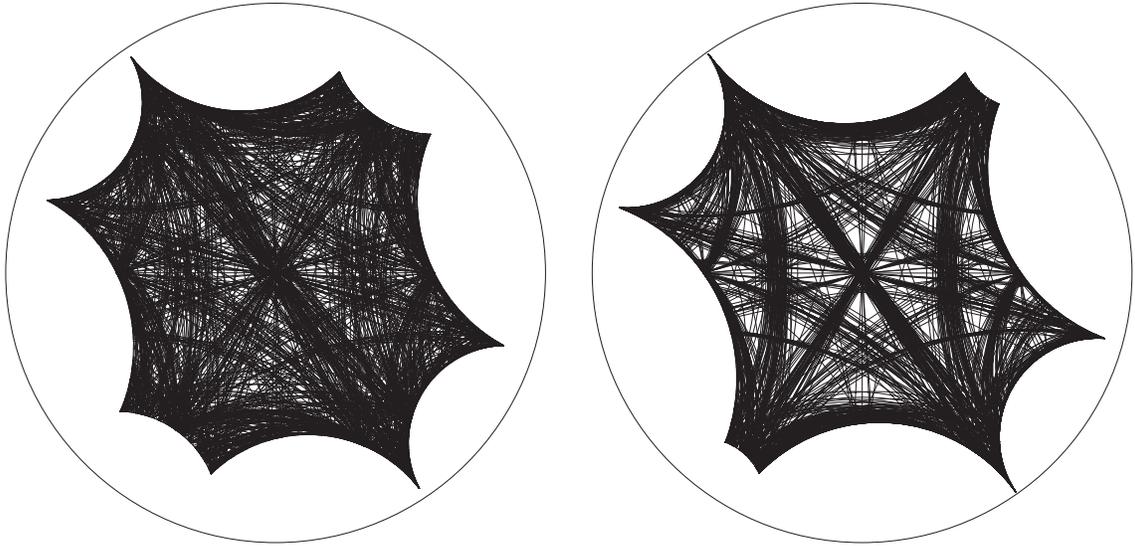


Figure 2: Perturbations of the Bolza surface

grow. Although we have no real evidence other than these figures, we wonder if the Bolza surface might be the surface where the largest gaps are the smallest.

Organization. The article is organized as follows. In Section 2 we prove Theorem 1.1, denoted Theorem 2.1 in the sequel. Section 3 is dedicated to the proof of Theorem 1.2, referred to later as Theorem 3.1. Thin surfaces and surfaces with cusps are treated in Section 4 where we prove Theorem 1.3, relabelled as Theorem 5.1. The article is concluded by an appendix which contains two technical results somewhat different in nature from the rest of the article.

Acknowledgements. We heartily thank Chris Judge, Manuel Racle, Klaus-Dieter Semmler, Juan Souto and Caroline Series for enlightening conversations and their encouragement.

2. GAPS ON THIN SURFACES

In this part we show the following result, where $a_g = \frac{1}{4 \cdot (4\pi(g-1))^2}$.

Theorem 2.1. *If $\text{sys}(S) \leq a_g$, then $C_S \geq a_g$.*

We shall show explicitly where on the surface a “forbidden disk” with the indicated radius may be found. For this we first construct a certain pair of pants on S .

Lemma 2.2. *Let γ be a simple closed geodesic on S of length $\ell(\gamma) < \frac{1}{2}$. Then there exists a pair of pants $Y \subset S$ with boundary geodesics $\gamma, \gamma_1, \gamma_2$ such that*

$$\cosh(\frac{1}{2}\ell(\gamma_i)) < \frac{\sinh(\frac{1}{2}\ell(\gamma))}{\ell(\gamma)} \cdot 4\pi(g-1), \quad i = 1, 2 \quad (1)$$

Before giving the proof we note that the length bound of $\frac{1}{2}$ on γ can be replaced by $\operatorname{arcsinh}(1)$ which is “optimal” for the argument we give.

Proof. We adapt an argument from [3, Section 5.2]. Cut S open along γ into a bordered surface \tilde{S} (consisting of either one or two connected components) with copies γ', γ'' of γ on the boundary. Take the component S' of \tilde{S} that has γ' on the boundary and look at the sets $C(r) = \{p \in S' \mid \operatorname{dist}(p, \gamma') \leq r\}$ for $r > 0$. For small r this set is an annulus and it does not intersect γ'' . As we let r grow, some first value r_γ will be reached where one of these two properties ceases to hold. Now, the injectivity radius near γ'' is far too small to allow the $C(r)$ to come close to γ'' as long as $r < r_\gamma$. (Their boundaries are simple closed curves of geodesic curvature smaller than the curvature of a horocycle). Hence, it's the annulus property that ceases to hold for $C(r_\gamma)$.

The rest is exactly as in [3, Section 5.2]: There are two geodesic segments of length r_γ emanating orthogonally from γ' and meeting each other under the angle π at their endpoints, thus forming a smooth geodesic arc η of length $2r_\gamma$. The endpoints of η on γ' dissect γ' into two arcs c_1, c_2 . The closed curves $c_1\eta$ and $\eta^{-1}c_2$ are freely homotopic to simple closed geodesics γ_1, γ_2 that together with γ' form the boundary of a pair of pants $Y \subset S'$ and by **formula [3, 2.3.4(i)]** the lengths satisfy $\cosh(\frac{1}{2}\ell(\gamma_i)) = \sinh(r_\gamma) \sinh(\frac{1}{2}c_i) < \sinh(r_\gamma) \sinh(\frac{1}{2}\ell(\gamma))$ for $i = 1, 2$. Since the interior of $C(r_\gamma)$ is still an annulus we have $\operatorname{area}(C(r_\gamma)) = \ell(\gamma) \sinh(r_\gamma) < \operatorname{area}(S) = 4\pi(g-1)$ and the lemma follows. \square

Throughout the paper we will be using hyperbolic trigonometry. As in the above proof, we will always refer to formula numbers from [3].

We also need an extension of the usual collar lemma [3, chapter 4] that includes complete non closed simple geodesics. For any simple closed geodesic γ on S the *width* is the quantity

$$w_\gamma = \operatorname{arcsinh}(1/\sinh(\frac{1}{2}\ell(\gamma))) \quad (2)$$

The collar theorem states among other things that the *collar* $\mathcal{C}_\gamma = \{p \in S \mid \operatorname{dist}(p, \gamma) < w_\gamma\}$ is homeomorphic to an annulus. (In contrast to the above $C(r)$ the collars \mathcal{C}_γ are defined as open sets.) The needed complement is the following.

Lemma 2.3 (Simple geodesics in collars). *Any complete simple geodesic intersecting \mathcal{C}_γ either intersects γ or converges to it.*

Proof. Let η be a complete simple geodesic on S , closed or non closed, that neither intersects γ nor converges to it. Let p be a point on η closest to γ . Let a be a simple geodesic arc from p to γ orthogonal to both η and γ and of length $\ell(a) = \text{dist}(\gamma, \eta) > 0$. In the universal cover of S there are lifts $\tilde{\gamma}$ of γ and $\tilde{\gamma}_1, \tilde{\gamma}_2$ of η , as in Figure 3 (which we use for two different purposes), together with lifts a_1, a_2 of a from $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ to $\tilde{\gamma}$ whose endpoints on $\tilde{\gamma}$ are distance $\ell(\gamma)$ apart from each other.

Since η is simple its lifts $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ are disjoint; they may have a common endpoint at infinity, though. We thus have a, possibly degenerated, right-angled geodesic hexagon with three consecutive sides of lengths $\ell(a), \ell(\gamma), \ell(a)$. This hexagon splits into two isometric pentagons. Applying **formula [3, 2.3.4(i)]** to either of them we get

$$\sinh(\ell(a)) \cdot \sinh\left(\frac{1}{2}\ell(\gamma)\right) \geq 1$$

and hence, $\text{dist}(\eta, \gamma) \geq w_\gamma$. □

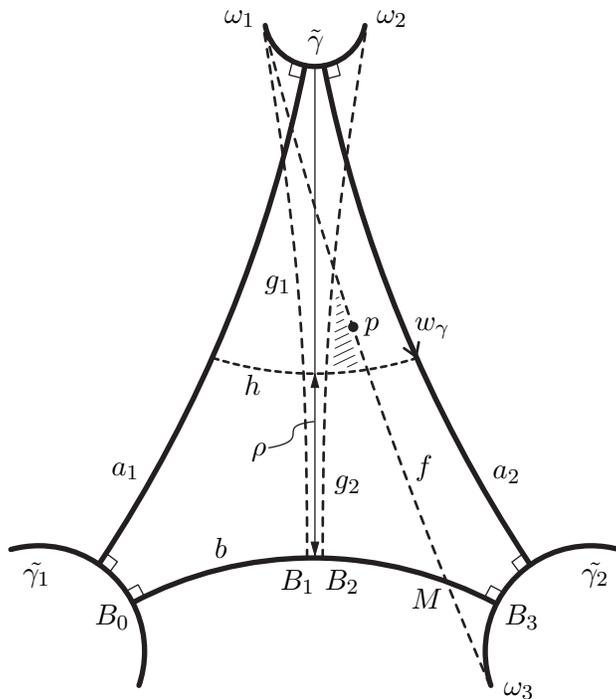


Figure 3: The geodesic hexagon G with a forbidden area

Proof of Theorem 2.1. Let γ be a simple closed geodesic of length $\ell(\gamma) \leq a_g$, where we think of a_g as being rather small; the bound for a_g as in the theorem will show up towards the end of the proof.

By Lemma 2.2 there exist simple closed geodesics γ_1, γ_2 with lengths bounded as in (1) that together with γ form the boundary of a pair of pants Y . We shall find a forbidden disk in the half collar $C(w_\gamma)$ in Y .

Y may be decomposed into two isometric right-angled geodesic hexagons by drawing the common perpendicular geodesic arcs a_1 from γ_1 to γ ; a_2 from γ to γ_2 ; and b from γ_2 to γ_1 . Figure 3 shows the lift G of one of the hexagons in the universal covering \mathbb{H} of S . (It need not be symmetric as drawn.) The lifts of the three perpendiculars are named after their originals, the lifts of the boundary geodesics are named, respectively, $\tilde{\gamma}_1, \tilde{\gamma}, \tilde{\gamma}_2$. We let ω_1, ω_2 be the endpoints at infinity of $\tilde{\gamma}$, labelled in such a way that ω_1 and a_2 lie on different sides of a_1 . We also consider the endpoint ω_3 at infinity of $\tilde{\gamma}_2$ that is separated from a_2 by b .

The dotted lines in Figure 3 are as follows. The vertical lines g_1, g_2 are the perpendicular geodesics from ω_1, ω_2 to b with respective endpoints B_1, B_2 on b . The horizontal line h consists of all points in the hexagon that have distance w_γ from $\tilde{\gamma}$. Line f is the geodesic from ω_1 to ω_3 .

We will prove below that if γ is as short as indicated, then the configuration of these lines is such that a shaded area Δ occurs, as drawn in the figure, consisting of points in the hexagon lying above h , to the right of g_2 and to the left of f . We will also estimate the size of Δ . Anticipating this for a moment we prove that Δ is forbidden, i.e., no lift of a simple complete geodesic of S in \mathbb{H} intersects Δ .

Since Δ lies in a lift of C_γ Lemma 2.3 implies that only then a simple geodesics η may have a lift passing through Δ when η intersect γ or converges to it. Now, take any such η and let $c \subset \eta$ be a connected component of $\eta \cap \text{interior}(Y)$ with one of its ends converging to γ (either by converging to a point on γ or by winding infinitely often around it). Let \tilde{c} be a component of its lift in \mathbb{H} that meets the part of G above line h . We have to show that $\tilde{c} \cap \Delta = \emptyset$. Since c has no self-intersections there are only the following three cases.

Case 1: the other end of c also converges to γ . Here we consider the geodesic symmetry $\sigma : \mathbb{H} \rightarrow \mathbb{H}$ with respect to the geodesic through b and observe that $G \cup \sigma(G)$ is a lift of Y . Hence, one end of \tilde{c} converges to $\tilde{\gamma}$ and the other to $\sigma(\tilde{\gamma})$. This implies that the part $\tilde{c} \cap G$ lies between g_1 and g_2 and cannot intersect Δ .

Case 2: the other end of c converges to γ_1 . In this case the other end of \tilde{c} converges to $\tilde{\gamma}_1$. This implies that, speaking with Figure 3, \tilde{c} lies on the left hand side of g_2 while Δ lies on the right hand side.

Case 3: the other end of c converges to γ_2 . By the same argument as before \tilde{c} now lies on

the right hand side of f and cannot intersect Δ either. This completes the proof that Δ is forbidden.

For the existence and the size of Δ we need a number of estimates beginning with the length of side $b = B_0B_3$. To that end, we abbreviate the right hand side of Inequality (1) by $\cosh \lambda$. The inequality in Lemma 2.2 then becomes $\frac{1}{2}\ell(\gamma_i) < \lambda$, ($i = 1, 2$). By **formula [3, 2.4.1(i)]** we have the following (where from now on we omit ℓ in the formulas):

$$\cosh\left(\frac{1}{2}\gamma\right) = \cosh(b) \sinh\left(\frac{1}{2}\gamma_1\right) \sinh\left(\frac{1}{2}\gamma_2\right) - \cosh\left(\frac{1}{2}\gamma_1\right) \cosh\left(\frac{1}{2}\gamma_2\right).$$

Here the right hand side is ≥ 1 and it increases when we replace $\frac{1}{2}\gamma_1$ and $\frac{1}{2}\gamma_2$ by λ . Using the identity $\cosh(b) \sinh^2(\lambda) - \cosh^2(\lambda) = 2 \sinh^2\left(\frac{1}{2}b\right) \sinh^2(\lambda) - 1$, we get

$$\sinh\left(\frac{1}{2}b\right) \sinh \lambda \geq 1. \quad (3)$$

The next estimate concerns the distance $\rho = \text{dist}(b, h)$, where we note that $\rho + w_\gamma$ is the length of the common perpendicular of b and $\tilde{\gamma}$. This perpendicular decomposes G into two right-angled pentagons. Applying **formula [3, 2.3.4(i)]** to the one that has the bigger side on $\tilde{\gamma}$ we get $\sinh(\rho + w_\gamma) \cdot \sinh\left(\frac{1}{4}\gamma\right) \leq \cosh \lambda$. Using that $e^\rho \sinh(w_\gamma) \leq \sinh(\rho + w_\gamma)$ and applying the definition (2) of w_γ we get the following, where $\tau = \sinh\left(\frac{1}{2}\gamma\right) / \sinh\left(\frac{1}{4}\gamma\right)$ is a factor close to 2:

$$e^\rho \leq \tau \cdot \cosh \lambda. \quad (4)$$

We now assume that the labelling of the sides of G has been set such that $\text{dist}(B_1, B_3) > \frac{1}{2}b$. Let M be the intersection point of b and f , where $f = \omega_1\omega_3$. Then M is the midpoint of the ideal crossed geodesic quadrilateral $\omega_1B_1B_3\omega_3$ and so we have an ideal right-angled triangle ω_1B_1M with side

$$B_1M > \frac{1}{4}b. \quad (5)$$

We estimate how far sides $g_1 = \omega_1B_1$ and ω_1M are apart from each other in the neighborhood of h . To this end, we take a point p on f and drop the perpendicular pp_1 to g_1 , the position of p being such that $\text{dist}(B_1, p_1) = \rho + r$ with $0 \leq r \leq \frac{1}{5}$. Applying **formula [3, 2.2.2(iv)]** to the ideal triangles ω_1B_1M and ω_1p_1p we get, using an obvious limit argument,

$$\frac{\tanh(pp_1)}{\tanh(B_1M)} = e^{-(\rho+r)}. \quad (6)$$

Finally, we provide a similar estimate for g_2 and g_1 . Applying the pentagon **formula [3, 2.3.4.(i)]** to half of the ideal quadrilateral $\omega_1B_1B_2\omega_2$ we have

$$\sinh\left(\frac{1}{2}B_1B_2\right) \cdot \sinh(\rho + w_\gamma) = 1.$$

Similarly to the preceding step we take a point q on g_2 and drop the perpendicular qq_1 to g_1 , the position of q being such that $\text{dist}(B_1, q_1) = \rho + r$ with $0 \leq r \leq \frac{1}{5}$. By **formula [3, 2.3.1(iv)]** and using that $\tanh(2t) \leq 2 \sinh(t)$, for $t \geq 0$, we have $\tanh(q_1q) = \cosh(\rho + r) \cdot \tanh(B_1B_2) \leq 2 \cosh(\rho + r) \sinh(\frac{1}{2}B_1B_2)$. Hence,

$$\tanh(q_1q) \leq 2 \frac{\cosh(\rho + r)}{\sinh(\rho + w_\gamma)} \leq 2 \frac{\cosh(r)}{\sinh(w_\gamma)} = 2 \cosh(r) \sinh(\frac{1}{2}\gamma). \quad (7)$$

By (7) on the one hand and by (3)–(6) on the other, by approximating the small terms in (3)–(7) linearly and using that $r \leq \frac{1}{5}$ so that $\cosh(r) < 1.03$ and $e^{-r} > 0.818$, we get:

$$\text{dist}(q_1, q) < 1.1 \ell(\gamma), \quad \text{dist}(p_1, p) > \frac{0.8}{(4\pi(g-1))^2}.$$

It follows that if $\ell(\gamma) \leq a_g$ with a_g as in the statement of the theorem, then in the vicinity of line h the points on g_2 are closer to g_1 than $1.1\ell(\gamma)$ and the points on f are further away than $3.2\ell(\gamma)$. Hence the shaded domain Δ shows up and one now easily sees that it contains a disk of radius a_g . \square

3. LOCAL QUANTIFICATION IN TERMS OF THE SYSTOLE

In this section S is a compact hyperbolic surface of genus g . We aim to show the following result.

Theorem 3.1. *Assume $s = \min\{\frac{1}{2}\text{sys}(S), \frac{1}{3}\}$. Then for any $\rho \leq s$ and any disk B_ρ of radius ρ in S there exists a point $p \in B_\rho$ such that*

$$\text{dist}(p, \mathcal{BS}) \geq \rho^2 e^{-M(g-1)},$$

where M is a constant that depends only on s .

We shall get the explicit bound $M = \frac{194}{s^2} \log(\frac{134}{s})$. Observe that Theorem 3.1 together with Theorem 2.1 provides a computable lower bound on the minimum value of C_S .

3.1. A Voronoi cell decomposition and its properties

We begin with a lemma that concerns ε -nets on S , by which we mean a set of points of pairwise distance at least ε and maximal for this property with respect to inclusion. The maximality of the set implies that the open balls of radius ε around the points cover the surface. Throughout we shall restrict ourselves to $\varepsilon \leq \frac{s}{2}$, where s is the systole of S .

Lemma 3.2. For fixed $0 < \varepsilon \leq \frac{s}{2}$ there exists an ε -net on S consisting of points $\{p_i\}_{i=1}^N$ with

$$N \leq \frac{2}{\cosh \varepsilon/2 - 1} (g - 1).$$

Proof. To construct such a set begin with an arbitrary point and add successively new points at least ε away from the preceding ones until this is no longer possible. The open balls of radius $\varepsilon/2$ around the resulting points p_i are pairwise disjoint, embedded and of area $2\pi(\cosh \varepsilon/2 - 1)$. The total area of the surface is $4\pi(g - 1)$ thus the number of balls cannot be greater than

$$\frac{4\pi(g - 1)}{2\pi(\cosh \varepsilon/2 - 1)}$$

which proves the lemma. \square

Given an ε -net one gets a cell decomposition of the surface given by the Voronoi cells associated to each point: each open cell V_i is the set of points whose closest point in the net is p_i . The boundary of V_i consist of points closest to more than one of the p_k s. By the maximality of the ε -net any boundary point of V_i is closer to p_i than ε . Since $\varepsilon \leq \frac{s}{2}$ it follows that V_i is contained in an embedded disc. Thus, the V_i are simple convex hyperbolic polygons. Since the vertices lie at distance $\leq \varepsilon$ from the center the sides of V_i have lengths $\leq 2\varepsilon$.

Remark 3.3. Note that up until now we haven't given any restrictions on how we choose the points. By standard perturbation techniques, if the points are chosen "generically", there will be exactly three cells adjacent to any vertex. By this we mean that by using a standard measure on the choice of points (using the measure on the surface), the choices where all vertices are *not* adjacent to exactly three cells lie in a measure 0 set. In all that follows, we shall suppose that this is the case. Hence, there will always be a triangulation dual to our Voronoi cell decompositions.

Lemma 3.4. Each Voronoi cell has at most $v(\varepsilon)$ sides, where $v(\varepsilon)$ is the integer part of

$$\pi / \operatorname{arccot} \left(\cosh(\varepsilon) \cdot \left\{ \sqrt{1 + 2 \cosh \varepsilon} + \sqrt{2 + 2 \cosh \varepsilon} \right\} \right).$$

Proof. Consider the triangulation dual to the Voronoi cell decomposition. Each side of a triangle has length at least ε . Any of the triangles is contained in a ball of radius ε (around the intersection point of the three Voronoi edges dual to the triangle). The number of sides of a Voronoi cell is the number of dual triangles that meet in its center. We now apply Lemma A.1 from the appendix. \square

From the two previous results we deduce a bound on the number of edges found in our Voronoi cell decomposition. We shall, however, assume from now on that ε is small. A restriction that has proved to be practical is $\varepsilon \leq \frac{1}{3}$.

Corollary 3.5. *Assuming that $\varepsilon \leq \min\{\frac{s}{2}, \frac{1}{3}\}$ we have $v(\varepsilon) = 12$, and there are at most*

$$12 \frac{g-1}{\cosh \varepsilon/2 - 1} < \left(\frac{97}{\varepsilon^2} - 10 \right) (g-1)$$

edges in a cell decomposition obtained as above.

Proof. The number of cells times the maximum number of edges per cell is an upper bound on twice the number of edges. Now $v(\varepsilon)$ from Lemma 3.4 is monotone increasing and $v(0) = v(\frac{1}{3}) = 12$. This proves the result. \square

In everything that follows in this section we shall suppose that we have a fixed Voronoi cell decomposition \mathcal{V} like the one we have just constructed with $\varepsilon \leq \min\{\frac{s}{2}, \frac{1}{3}\}$.

3.2. Scheme of proof

In this subsection we describe the strategy of our proof without any computations in order to motivate the estimates we perform afterwards.

Given an embedded disk $D = B_\rho$ of radius $\rho > 0$ on S , we look at the restriction of the Birman-Series set to D . This consists of a countable collection of geodesic segments between points on the boundary of D .

To find quantifiable “empty space” between these segments we begin by introducing a constant $R > 0$ that later will be adjusted. For each segment c we look at a larger geodesic segment \bar{c} obtained by extending c outside of D . Specifically, we take the continuation of c of length $R - 2\varepsilon$ in both directions (measured from the midpoint of c). This larger “segment” is not necessarily embedded but as c belongs to the Birman-Series set, it does not contain any transversal self-intersections. Either endpoint of \bar{c} lies in some Voronoi cell of \mathcal{V} . (The two cells may coincide.)

To “capture” \bar{c} we associate to it what we shall call a *model strand* m^c the precise construction of which shall follow in the next subsection. Among its properties we have the following. The endpoints of m^c lie on vertices of the 1-skeleton of \mathcal{V} and m^c is freely homotopic to \bar{c} , where one allows endpoints to move within the start and terminal Voronoi cells of \bar{c} . The important part of m^c is its middle intersection with D , i.e., the connected component m_c of $m^c \cap D$ that contains the midpoint of m^c . We call m_c a *model strand in* D .

The model strands shall allow us to quantify the arguments in [2] by showing that the number of all m_c grows polynomially with R while the distance between c and m_c has an upper bound that decays exponentially with R . As a result, for large enough R there are “relatively few” model strands in D and each segment c lies in an “extremely small” tubular neighborhood of some of these. Outside these tubular neighborhoods we shall then have empty space whose size can be estimated in terms of R .

The model strands will be obtained in two steps. First we associate to \bar{c} a combinatorial path $P(\bar{c})$ on the 1-skeleton of \mathcal{V} that is simple in the combinatorial sense (see the next subsection). The number of such paths can be estimated using combinatorial arguments. In the second step we then homotope the path $P(\bar{c})$ into a geodesic arc m^c keeping the endpoints fixed. It may turn out that m^c has self-intersections, but m_c is simple. The distance between c and m_c is estimated in Section 3.5.

3.3. Constructing combinatorial arcs and their properties

Here we describe a method for associating to any simple arc such as \bar{c} a combinatorial path that lies in the 1-skeleton of \mathcal{V} .

Consider a finite oriented geodesic arc a on S which we assume to be simple in the sense that it has no transversal self-intersections. We also assume that a is longer than 4ε (in our applications a will be much larger). Then a begins in some Voronoi cell V_0 , traverses a sequence of cells $\{V_i\}_{i=1}^n$ and ends up in some cell V_{n+1} (each cell may appear multiple times). We break up a into smaller arcs a_i where for $i = 0, \dots, n+1$ each a_i is a connected component of $a \cap V_i$ and a_i is connected to a_{i+1} ($i \leq n$). Intersections that consist of a vertex only are ignored, that is we list the cells such that each a_i has positive length. For the combinatorial path we proceed as follows.

- a_0 and a_{n+1} do not contribute to our combinatorial path.
- For $i = 1, \dots, n$, if a_i intersects the interior of V_i we associate to it the homotopic polygonal arc P_i on the boundary ∂V_i that has the same endpoints as a_i and uses the minimal number of edges of V_i (see Figure 5).
- If a_i is a side of V_i we set P_i to be a_i .

We add a refinement to item 2: if V_i has an even number of sides, say m_i , and if a_i connects a pair of opposite sides of V_i , then there are two choices for the path P_i , either using $\frac{1}{2}m_i + 1$ edges of V_i . To remove this ambiguity we choose, for any Voronoi cell V , a “separator point”

p_V (not necessarily the geometric center of V) in the interior of $V \setminus a$ and then require that in the above cases P_i is the path homotopic to a_i in the punctured disk $V_i \setminus \{p_{V_i}\}$.

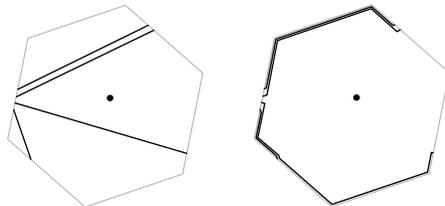


Figure 4: Pushing the path to the 1-skeleton

The concatenation of these paths, $P_1 \cup \dots \cup P_n$, yields a connected path that lies entirely on the 1-skeleton of \mathcal{V} , possibly with partial edges. We shall remove the latter by “shrinking” homotopies that take place on the edges of the 1-skeleton as indicated in Figure 6 so as to obtain a purely combinatorial connected sequence of edges $P(a)$ of the Voronoi polygons.

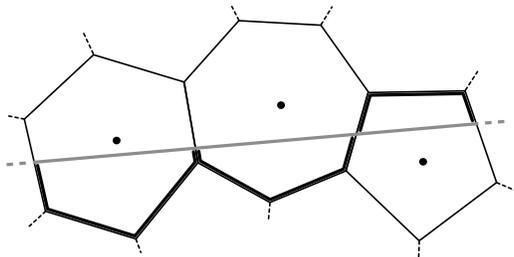


Figure 5: Obtaining a polygonal path

Thus, the construction is in two steps, the first step consisting in “pushing the arcs to the boundary” as indicated in Figure 4, the second step being the shrinking away of the partial edges.

The combinatorial path $P(a)$ obtained in this way is simple in that it does not contain any transversal self-intersections. To see this one may consider an arbitrarily thin tubular neighborhood T around the 1-skeleton and slightly modify the pushing and shrinking homotopies so as to obtain a genuine simple path $P'(a)$ in T with the same combinatorics as $P(a)$.

We define the *combinatorial length* of $P(a)$ to be the number of edges it contains. The next lemma provides an upper bound of this length in terms of the original path length.

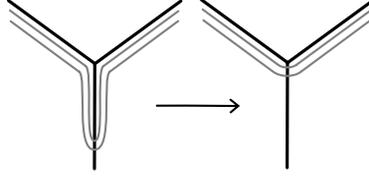


Figure 6: Elimination of partial edges by shrinking

Lemma 3.6. *A path a as above of length ℓ passes through at most*

$$\frac{4}{\varepsilon}(\ell + 3\varepsilon)$$

Voronoi cells. This multiplied by 6 is an upper bound for the combinatorial length of $P(a)$.

Proof. We argue in the universal cover by considering a lift \tilde{a} of a . Along \tilde{a} we have the sequence of lifts $\tilde{V}_1, \dots, \tilde{V}_n$ of the Voronoi cells V_1, \dots, V_n with the property that $\tilde{V}_i \cap \tilde{a}$ is a lift of a_i , $i = 1, \dots, n$ (recall that a_0 and a_{n+1} are not taken into account). The interiors of the lifted cells are pairwise disjoint and so are the balls of radius $\varepsilon/2$ around their centers q_1, \dots, q_n . Furthermore, these balls lie in a tubular neighborhood of radius $3\varepsilon/2$ around \tilde{a} . We now apply an area argument. The total area of the $\varepsilon/2$ -balls around the points $\{q_i\}_{i=1}^n$ is

$$n \cdot 2\pi(\cosh \varepsilon/2 - 1).$$

The area of the $3\varepsilon/2$ neighborhood of \tilde{a} is

$$2(\ell \sinh 3\varepsilon/2 + \pi(\cosh 3\varepsilon/2 - 1)).$$

Area comparison and elementary simplification using that $\varepsilon \leq \frac{1}{3}$ now yields the first statement. The second statement follows from Lemma 3.4 using that $v(\varepsilon) \leq v(\frac{1}{3}) = 12$ and that for the crossing of any cell V_i the combinatorial path $P(a)$ runs along at most half the edges of V_i . \square

Remark 3.7. Although this is not needed in what follows we indicate a reverse inequality giving a lower bound on the number of Voronoi cells a path a of length ℓ as in the lemma necessarily traverses. To see this consider a maximal set of points 2ε apart on the covering \tilde{a} . There are at least $\ell/2\varepsilon$ such points. As Voronoi cells have diameters less than 2ε this means that \tilde{a} , respectively a needs to traverse at least

$$\frac{\ell}{2\varepsilon} - 1$$

Voronoi cells. There is also a lower bound on the combinatorial length L of $P(a)$: using that the edges of a Voronoi cell are not longer than the diameter the triangle inequality yields $2\varepsilon(L + 2) \geq \ell$ and thus

$$L \geq \frac{\ell}{2\varepsilon} - 2.$$

3.4. Counting combinatorial paths

We need to bound the number of combinatorial paths in terms of the lengths of the model strands. Using the lemma above, this can be obtained via the combinatorial lengths. For the following we denote by E the number of edges of our Voronoi decomposition \mathcal{V} .

Lemma 3.8. *There are at most $4L^2 \binom{L+E}{L}$ combinatorial simple paths of length at most L .*

Proof. The model for the argument that follows is the following useful fact: the homotopy class of a (simple) multicurve lying on a triangulated surface is determined by its intersection numbers with each of the sides of the triangulation. The triangulation in our case is the triangulation dual to the Voronoi cells and the intersection numbers are exactly the numbers of times the path traverses a given edge. So if our path was closed it would be uniquely determined by the number of times it passes through each edge.

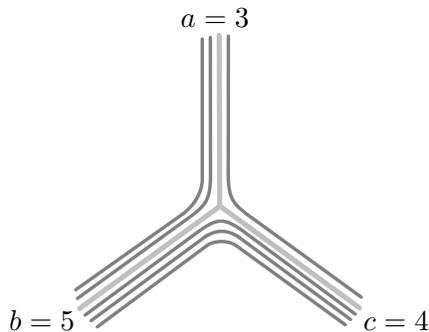


Figure 7: Example of how the numbers of parallel segments determine the combinatorial paths

We now modify this for non closed paths. In order not to confuse path edges with the E edges of the 1-skeleton of the Voronoi decomposition, we shall call path edges *segments*.

We begin by distributing a number $\leq L$ of segments (later to be concatenated) among the E edges. By elementary combinatorics this is possible in $\binom{L+E}{L}$ different ways. For ease of description we place them as distinct segments parallel to the corresponding edges in a thin tubular neighborhood of the 1-skeleton of \mathcal{V} as drawn in Figures 7 and 8.

Next we select, for any such distribution, a pair of segments that shall play the role of the end segments of the path to be constructed and erase half of each of the two segments. This is possible in at most $4L^2$ different ways. Figure 8 shows an example with half a segment erased.

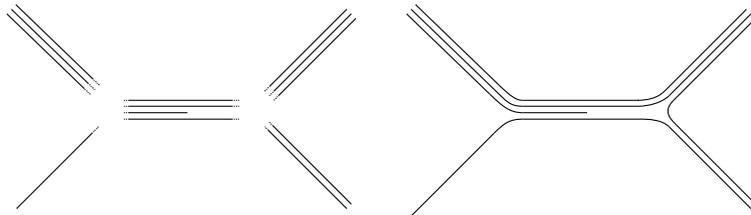


Figure 8: An example of how to construct the path with a given starting segment

At any vertex of the 1-skeleton there are now a number of incoming segments and there is at most one way to paste these together at the endpoints such that the resulting arcs do not intersect each other (pairs of segments on the same edge are not allowed to be pasted together). Of course, not every distribution allows one to paste all segments at all vertices, and even if the pasting is possible the result may be disconnected; but every simple path of length $\leq L$ may be obtained by some distribution and, hence, there are at most $4L^2 \binom{L+E}{L}$ such paths. \square

3.5. Bounding the distance between arcs and model arcs

In what follows, we need to metrically compare the model arcs to the arcs they are intended to approximate. This is achieved via the following lemma about arcs in the hyperbolic plane.

Lemma 3.9. *Fix $\delta > 0$, $\rho > 0$. In the hyperbolic plane, let b, b' be two geodesic arcs of lengths $> 2(\rho + \delta)$ such that the two initial points and the two endpoints are at respective distances $\leq \delta$ from each other. Furthermore, consider a disk D of radius ρ centered at the midpoint of b . Then $D \cap b$ lies in an r -neighborhood of b' , where*

$$r \leq \operatorname{arcsinh} \left(\frac{\cosh \rho \sinh \delta}{\cosh \frac{\ell(b)}{2}} \right).$$

Proof. We search for an “extremal” b' that shall allow us to compute the constants appearing in a worst case scenario. For this we consider the two disks of radius δ surrounding the endpoints of b .

An extremal b' must have its endpoints on the boundary of these disks and a moment’s reflection shows that the worst case scenario is given by the two geodesics tangent to the

boundary of these disks that do not cross b (see Figure 9).

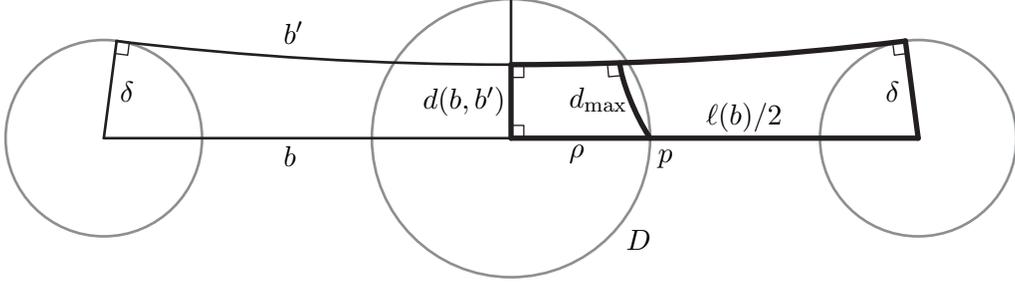


Figure 9: Comparing the two arcs

Let b' be one of them. By symmetry, the distance path between b and b' reaches b at its midpoint. One can now compute in the trirectangles as shown in the figure. The bigger trirectangle has sides of lengths δ , $\frac{1}{2}\ell(b')$, $d(b, b')$, $\frac{1}{2}\ell(b)$. Using hyperbolic trigonometry formula [3, 2.3.1(v)] we obtain

$$\sinh \delta = \sinh d(b, b') \cosh \frac{\ell(b)}{2}.$$

The least upper bound d_{\max} for the distances to b' of points on $D \cap b$ is reached for the point p where b intersects the boundary of D . It remains to show that for $r := d_{\max}$ the inequality as in the lemma is satisfied. We actually have equality: p is the vertex of the smaller trirectangle with sides d_{\max} , ρ , $d(b, b')$, ρ ; using the same formula as before we obtain

$$\sinh d_{\max} = \sinh d(b, b') \cosh \rho$$

and putting the two formulas together we get

$$\sinh d_{\max} = \frac{\cosh \rho \sinh \delta}{\cosh \frac{\ell(b)}{2}}.$$

□

3.6. Estimates and finalizing the proof

We first collect and simplify some of the earlier bounds.

Number of edges. In view of Corollary 3.5, we use the abbreviations

$$G := \frac{97(g-1)}{\varepsilon^2}, \quad G' := G - 10. \quad (8)$$

By the same corollary, \mathcal{V} has at most G' edges.

Combinatorial length. As described in Section 3.2 any arc c in the disk $D = B_\rho$ that occurs as a connected component of $D \cap \gamma$ for some simple complete geodesic γ on S is extended to a larger arc \bar{c} on γ . The extension goes in both directions, starting from the midpoint of c , each half of the extension having length $R - 2\varepsilon$. To \bar{c} we associate the combinatorial path $P(\bar{c})$ as described in Section 3.3 (with \bar{c} in the role of a). The extension \bar{c} has length $\ell = 2R - 4\varepsilon$ and by Lemma 3.6 the combinatorial length of $P(\bar{c})$ is bounded above by

$$L_R := \frac{48}{\varepsilon}R. \quad (9)$$

The number of model strands. By Lemma 3.8 there are at most

$$4L^2 \binom{L+G'}{L} \leq \frac{4L^2(L+G')^{G'}}{G'!} \leq \frac{4G^2(L+G)^G}{G!} \quad (10)$$

combinatorial paths $P(\bar{c})$ of combinatorial length L . For $L = L_R$ this is at the same time an upper bound for the number of model strands m^c , respectively the number of strands m_c in D .

A heuristic check shows that the area argument that will follow further down can only succeed if

$$R > G.$$

We shall therefore work from now on under this hypothesis. Using Stirling's formula and the fact that $G \geq \frac{97}{\varepsilon^2}$ we then get the following bound for the number of model strands which has been tailored in view of its later application:

$$N(R) := \frac{1}{10} \frac{m^G}{G^G} R^G \quad \text{with} \quad m = \frac{134}{\varepsilon}. \quad (11)$$

Distance between c and m_c . By construction the two initial, respectively endpoints of \bar{c} and its associated model strand m^c lie in the same Voronoi cells and their respective distances are smaller than 2ε . Furthermore, \bar{c} has length $2R - 4\varepsilon$. By Lemma 3.9, c lies in an r -neighborhood of m_c where, by elementary simplification and using that $\rho \leq \varepsilon \leq \frac{1}{3}$,

$$r \leq \frac{\cosh \rho \sinh 2\varepsilon}{\cosh(R - 2\varepsilon)} \leq 3e^{-R} =: w_R. \quad (12)$$

Area argument. This is the heart of the proof. For subsets $A \subset D$ and $t > 0$ we shall denote by A^t the part of the t -neighborhood of A that lies in D . For each model strand m_c in D the set $m_c^{2w_R}$ has area

$$\text{area}(m_c^{2w_R}) \leq 2\ell(m_c) \sinh(2w_R) \leq 4\rho \sinh(2w_R) < 9\rho w_R. \quad (13)$$

Now let \mathcal{M} be the union of all model strands in D . Then \mathcal{M}^{2w_R} has area less than $9\rho w_R$ times the number of model strands in D , while D has area $2\pi(\cosh(\rho) - 1) > \pi\rho^2$. On the other hand \mathcal{BS} is contained in \mathcal{M}^{w_R} . Hence, if we can determine R such that the area bound for \mathcal{M}^{2w_R} is smaller than $\pi\rho^2$, then \mathcal{M}^{2w_R} does not cover D . By (12) we therefore get a point $p \in D$ at distance w_R from \mathcal{BS} if we take R to be a solution $> G$ to the equation

$$\frac{1}{\rho} \frac{m^G}{G^G} R^G = e^R. \quad (14)$$

Estimating the solution and end of proof. We first state the following lemma which shall also be used in the next section.

Lemma 3.10. *Let α, γ be positive constants, $\alpha > \frac{e^\gamma}{\gamma^\gamma}$. Then the equation $\alpha t^\gamma = e^t$ for $t > 0$ has two solutions $t_1 < \gamma, t_2 > \gamma$ and*

$$\log(\alpha\gamma^\gamma) < t_2 < 2\log(\alpha\gamma^\gamma).$$

Proof. The equation is equivalent to $\gamma\alpha^{\frac{1}{\gamma}} \frac{t}{\gamma} = e^{\frac{t}{\gamma}}$. Substituting $\tau = \frac{t}{\gamma}$ we transform it into

$$\frac{e^\tau}{\tau} = \beta \quad (15)$$

for $\tau > 0$ with $\beta = \gamma\alpha^{\frac{1}{\gamma}}$. By the hypothesis on α we have $\beta > e$ and so (15) has two solutions $\tau_1 < 1$ and $\tau_2 > 1$. For $\tau' = \log \beta$ we have $\tau' > 1$ and $\frac{e^{\tau'}}{\tau'} < \beta$. For $\tau'' = 2\log \beta$ we have $\tau'' < \beta$ and $\frac{e^{\tau''}}{\tau''} > \beta$. Hence, $\log \beta < \tau_2 < 2\log \beta$. Substituting back we get the claims of the lemma. \square

In the case of equation (14) we have $\gamma = G$ and $\alpha = \frac{1}{\rho} \frac{m^G}{G^G}$. By Lemma 3.10 the larger of the two solutions to the equation has the bound $R \leq 2\log(\frac{1}{\rho}) + 2G\log(m)$. We thus get the lower bound on the maximal distance to \mathcal{BS} in D :

$$w_R \geq 3\rho^2 e^{-2G\log m}. \quad (16)$$

For $\varepsilon = s$ this yields the bound as stated in Theorem 3.1.

4. SURFACES WITH CUSPS AND (VERY) SMALL GEODESICS

In this section we generalize the local quantification results of the previous sections to include surfaces with cusps and small geodesics. The process will be almost identical to previously with just a number of necessary changes that will be detailed. Again ε is a fixed constant, $0 < \varepsilon \leq \frac{1}{3}$.

4.1. Thick and thin decomposition

We consider this time a complete orientable finite area hyperbolic surface S of genus g with $n \geq 0$ cusps. A closed geodesic on S shall be called *small* if its length is $\leq 2\varepsilon$. We let β_1, \dots, β_h be the list of all small geodesics of S arranged with decreasing lengths

$$2\varepsilon \geq \ell(\beta_1) \geq \ell(\beta_2) \geq \dots \geq \ell(\beta_h).$$

The list may be void, but we shall assume that $n + h > 0$ to contrast this section from the preceding ones. We make use of the *collar theorems* as e.g. Theorems 4.1.1, 4.1.6 and 4.4.6 in [3] of which we recall the following.

For the cusps we have the *cuspidal neighborhoods* \mathcal{P}_i , $i = 1, \dots, n$, filled by the horocycles of lengths < 2 . Each β_k is simple and has a *collar neighborhood* \mathcal{C}_k , $k = 1, \dots, h$, filled by the points at distance $< w(\beta_k)$ from β_k , where $\sinh(w(\beta_k)) \sinh(\frac{1}{2}\ell(\beta_k)) = 1$. The two boundary curves of \mathcal{C}_k are parallel curves to β_k (i.e. all points have the same distance $o(\beta_k)$) and their lengths satisfy $\ell(\beta_k) \cdot \cosh(w(\beta_k)) > 2$. Topologically the \mathcal{P}_i are punctured discs, the \mathcal{C}_k are annuli and all these neighborhoods are pairwise disjoint. Finally we note that

$$h \leq 3g - 3 + n. \tag{17}$$

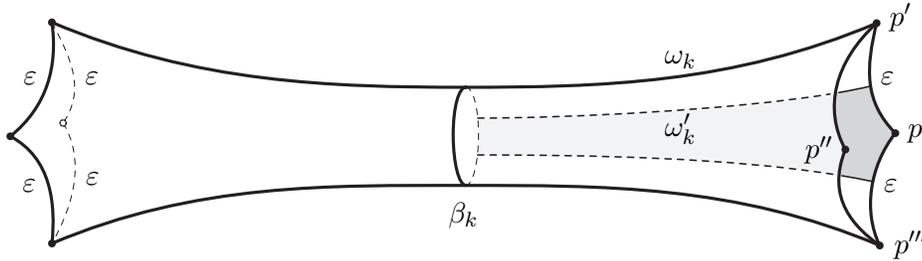


Figure 10: Reduced collar \mathcal{C}'_k with part of the Voronoi cell at vertex p

For $k = 1, \dots, h$ we now choose $\omega_k < w(\beta_k)$ in such a way that on either side of β_k the parallel curve at distance ω_k admits a quadruple of points at consecutive distances ε as the points p, p', p'', p''' shown in Figure 10. The value of ω_k is given by

$$\sinh(\varepsilon/2) = \sinh\left(\frac{1}{8}\ell(\beta_k)\right) \cosh(\omega_k),$$

which is **formula [3, 2.3.1(v)]** applied to the trirectangle with acute angle at p' and consecutive sides of lengths $\frac{\varepsilon}{2}$, ω_k , $\frac{1}{8}\ell(\beta_k)$, ω'_k , where ω'_k is the distance between β_k and the segment from p to p' . From the formulas for $w(\beta_k)$ and ω_k we deduce by elementary

computation that indeed $\omega_k < w(\beta_k)$. More accurately (using a numerical check for the upper bound),

$$\operatorname{arccosh}(2) < \omega_k < w(\beta_k) - \frac{1}{3}. \quad (18)$$

The same formula also yields the injectivity radius r_p at p (and p' , etc.). Indeed, the shortest geodesic loop at p together with β_k forms two identical trirectangles with consecutive sides $\omega_k, \frac{1}{2}\ell(\beta_k), \cdot, r_p$ and we have $\sinh(r_p) = \sinh(\frac{1}{2}\ell(\beta_k)) \cosh(\omega_k)$. Bringing the two formulas together we obtain

$$\sinh(r_p) = \frac{\sinh(\frac{1}{2}\ell(\beta_k))}{\sinh(\frac{1}{8}\ell(\beta_k))} \sinh(\varepsilon/2) > 4 \sinh(\varepsilon/2).$$

Since $\varepsilon \leq \frac{1}{3}$ we have $r_p > 1.8\varepsilon$.

The geodesic segments of length ε that connect the points p, p', p'', p''', p form a simple curve homotopic to β_k and there is another such curve on the other side of β_k . We choose it in such a way that the subset \mathcal{C}'_k of \mathcal{C}_k that lies between these two curves is symmetric with respect to β_k as shown in Figure 10. We call \mathcal{C}'_k the *reduced collar*.

In a similar manner we define the reduced cusp neighborhoods $\mathcal{P}'_i \subset \mathcal{P}_i$, $i = 1, \dots, n$. For each \mathcal{P}_i the vertices p, p', p'', p''' on the boundary of \mathcal{P}'_i lie on a horocycle. Since cusps may be viewed as limits of half collars we may apply the previous estimates taking the limit for $\ell(\beta_k) \rightarrow 0$. In particular the injectivity radius at the points p, p' , etc. has again the lower bound $r_p > 1.8\varepsilon$. We call the union of the reduced cusps and collars the thin part of S and the complement

$$S' = S \setminus (\mathcal{C}'_1 \cup \dots \cup \mathcal{C}'_h \cup \mathcal{P}'_1 \cup \dots \cup \mathcal{P}'_n)$$

the thick part. By the collar theorems and the lower bound on r_p the injectivity radius at any point in S' is larger than ε .

The main result is the following.

Theorem 4.1. *Let S be as described and set*

$$\sigma = \ell(\beta_1) \cdots \ell(\beta_h)$$

if $h \geq 1$ and $\sigma = 1$ if $h = 0$. Then for any $\rho \leq \varepsilon$ and any disk B_ρ of radius ρ in the thick part of S there exists a point $p \in B_\rho$ such that

$$\operatorname{dist}(p, \mathcal{BS}) \geq \rho^2 \sigma^2 e^{-M(2g-2+n)},$$

where M is a constant that depends only on ε .

The proof will be finalized in section 4.6, where we shall get the explicit bound $M = \frac{97}{\varepsilon^2} \log(\frac{134}{\varepsilon})$.

4.2. An ε -net in the thick part and Voronoi cells in S

We construct an ε -net in S' beginning with the above vertices p, p', \dots , on the boundaries of the reduced collars and cusp neighborhoods. We call these points the “special points”. By the estimates in the preceding section and because collars and cusps are pairwise disjoint the special points have pairwise distances $\geq \varepsilon$. We now complete them into an ε -net on S' by successively adding additional “ordinary points” on S' at pairwise distances $\geq \varepsilon$ until this is no longer possible. For the properties of the resulting Voronoi cells we first prove the following.

Lemma 4.2. *For $x \in S \setminus S'$ the distance to the special points of the ε -net is smaller than the distance to the ordinary points.*

Proof. Let q be an ordinary point of the ε -net. The shortest connection from x to q intersects one of the boundary segments of length ε in some point y . One of the endpoints of the segment, say p , satisfies $\text{dist}(y, p) \leq \varepsilon/2$. We now have $\text{dist}(x, p) \leq \text{dist}(x, y) + \text{dist}(y, p) \leq \text{dist}(x, y) + \varepsilon/2$, where at least one of the inequalities is strict. On the other hand $\text{dist}(x, q) = \text{dist}(x, y) + \text{dist}(y, q) \geq \text{dist}(x, y) + \text{dist}(p, q) - \varepsilon/2 \geq \text{dist}(x, y) + \varepsilon/2$. Altogether $\text{dist}(x, p) < \text{dist}(x, q)$ which proves the claim. \square

Our ε -net will now be considered as a distribution of points on S and defines a decomposition of S into Voronoi cells. We shall call “special Voronoi cells” those around special points and “ordinary Voronoi cells” those around ordinary points. The cell decomposition consisting of all special and all ordinary Voronoi cells is again denoted by \mathcal{V} . The following properties take over from those in Section 3.1:

Lemma 4.3. *Any cell of \mathcal{V} contains an embedded disk of radius $\varepsilon/2$ and has at most 12 sides.*

Proof. Lemma 4.2 implies that the ordinary Voronoi cells are contained in S' and have the same properties as those in Section 3.1. For the special cells we argue as follows. Consider, for instance, the special point p on the boundary of the reduced collar \mathcal{C}'_k as shown in Figure 10. It follows from Lemma 4.2 that the part in \mathcal{C}'_k of the Voronoi cell at p is the shaded polygon shown in the figure with two sides of length $\varepsilon/2$ on the segments adjacent to p , then the two perpendiculars from these segments to β_k and finally a side of length $\frac{1}{4}\ell(\beta_k)$ on β_k . The domain may be decomposed into two identical trirectangles with acute angle, say φ , at p . **Formula [3, 2.3.1(iii)]** yields $\cosh(\frac{1}{8}\ell(\beta_k)) = \cosh(\varepsilon/2) \sin \varphi$. Given that $\varepsilon \leq \frac{1}{3}$ and $\ell(\beta_k) \leq 2\varepsilon$ we deduce from it that the obtuse angle of the shaded domain

at p is $2\varphi \geq 0.9\pi$. Proceeding as in the proof of Lemma 3.4 we now conclude that the Voronoi cell at p has at most 9 sides. (From a combinatorial point of view, however, we count an additional degenerate side on β_k since at the endpoints of the sides on β_k there are four meeting cells.) Finally, with a glance at Figure 10 we see that the cell contains an embedded disk of radius $\varepsilon/2$ centered at p . \square

When p is on the boundary of a reduced cusp neighborhood then the result is the same except that the shaded domain is degenerate and one side is replaced by a point at infinity. Summing up we have the same conclusion as in Corollary 3.5 using that the area of S is now $2\pi(2g - 2 + n)$

Corollary 4.4. *The number of sides of \mathcal{V} is bounded above by $(\frac{97}{\varepsilon^2} - 10)(g - 1 + \frac{n}{2})$.*

4.3. Traversing and terminal arcs

We now proceed as in the previous section to construct our model strands. As before, we consider an embedded disk D of radius $\rho > 0$ and a geodesic segment c in D belonging to the Birman-Series set. This time, however, D is contained in the thick part S' . We take the continuation \bar{c} of c in both directions of length $R - 2\varepsilon$. The endpoints lie in Voronoi cells. The only difference with the previous sections is the way in which we deal with the special Voronoi cells. The intersections of \bar{c} with the collars and cusps may be of two kinds: *traversing arcs* and *terminal arcs*. The precise definition and properties are as follows.

Traversing arcs. Figure 11 shows half of a traversing arc \mathcal{T} lifted to the universal covering of S . The horizontal line $\tilde{\beta}_k$ on the top is a lift of the geodesic β_k , the lower end is a lift of one of the boundary curves of \mathcal{C}'_k consisting of geodesic segments of length ε . The dotted line is tangent to these and is a curve parallel to $\tilde{\beta}_k$ at distance ω'_k . The vertices have distance ω_k to $\tilde{\beta}_k$. The meanings of ω_k and ω'_k are the same as in Section 4.1 (Figure 10).

The projection (under the universal covering map) on S of the strip between the dotted lines is the *dotted collar* $\mathcal{C}''_k \subset \mathcal{C}'_k \subset \mathcal{C}_k$ defined as

$$\mathcal{C}''_k = \{x \in S \mid \text{dist}(x, \beta_k) \leq \omega'_k\}. \quad (19)$$

The boundary curves of \mathcal{C}''_k are tangent to the boundary curves of \mathcal{C}'_k . In a similar way we define (for later use) in each cusp neighborhood \mathcal{P}_i the *dotted cusp neighborhood* $\mathcal{P}''_i \subset \mathcal{P}'_i \subset \mathcal{P}_i$ whose boundary curve is a horocycle tangent to the boundary curve of \mathcal{P}'_i .

Note that the traversing arcs begin and end on the dotted lines, i.e., any traversing arc \mathcal{T}

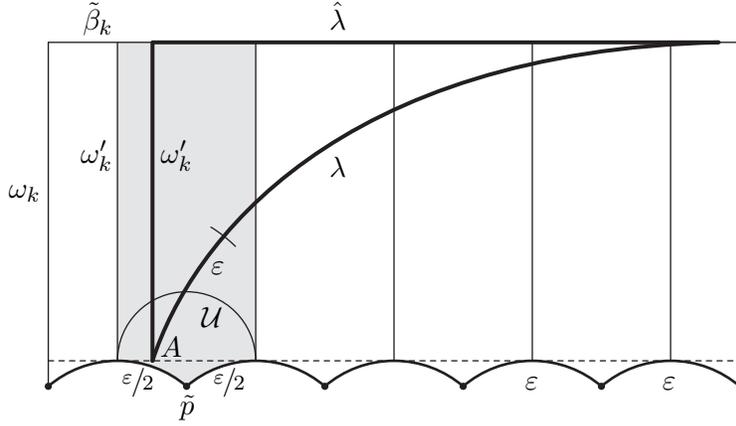


Figure 11: Traversing arc depicted in the universal covering

is a connected component, for some $k = 1, \dots, h$, of $\bar{c} \cap \mathcal{C}_k''$ that has its endpoints on the two opposite boundary curves of \mathcal{C}_k'' .

The arc labelled λ in Figure 11 is half of the lift of \mathcal{T} going from the dotted line to $\tilde{\beta}_k$. The label also denotes the length of this arc. Projecting λ orthogonally to $\tilde{\beta}_k$ we obtain the leg $\hat{\lambda}$ of a right-angled geodesic triangle with hypotenuse λ whose other leg is ω'_k . We now compare λ with $\hat{\lambda}$.

Formula [3, 2.2.2(i)] applied to the aforementioned right angled triangle yields $\cosh(\lambda) = \cosh(\hat{\lambda}) \cosh(\omega'_k)$, and by **formula [3, 2.3.1(iv)]** applied to the trirectangle with sides $\frac{1}{2}\varepsilon, \omega'_k, \frac{1}{8}\ell(\beta_k), \omega_k$ we have $\cosh(\omega'_k) = \tanh(\varepsilon/2) \coth(\frac{1}{8}\ell(\beta_k))$. Bringing this together and recalling that $\ell(\beta_k) \leq 2\varepsilon, \varepsilon \leq 1/3$ we get $\cosh(\lambda) \geq \cosh(\hat{\lambda}) \tanh(\varepsilon/2) \coth(\varepsilon/4) \geq 1.98 \cosh(\hat{\lambda})$ and then, by an elementary estimate,

$$\lambda - \hat{\lambda} > 2/3. \quad (20)$$

The endpoint A of λ on the dotted line lies in a disk \mathcal{U} of radius $\varepsilon/2$ around one of the vertices \tilde{p} . By the triangle inequality this implies, in turn, that \mathcal{U} is contained in the disk of radius ε around A .

Terminal arcs. These are the connected components of the intersections of \bar{c} with the dotted collars and cusp neighborhoods that have one endpoint on the boundary while the other lies in the interior. There are at most two such components.

Since the infinite geodesic extension of \bar{c} is a simple curve it follows that terminal arcs in a cusp neighborhood are “vertical” that is, orthogonal to the horocycles. In the collars the situation is different.

To deal with the terminal arcs in \mathcal{C}_k ($k = 1, \dots, h$) we supplement the 1-skeleton of \mathcal{V} by a

Some of the comparison arcs will go from \tilde{q} to these points, but we need further ones. To this end we look at the geodesic ray η emanating from \tilde{q} asymptotic to $\tilde{\beta}_k$ as shown in Figure 12. Together with $\tilde{\beta}_k$ and the vertical line at \tilde{q} it forms an ideal right-angled triangle with finite leg ω'_k and acute angle at \tilde{q} which we write as a sum $\phi + \theta$, ϕ being the angle at \tilde{q} of the right-angled triangle with hypotenuse $R'/2$ and legs ω'_k, t . By **formula [3, 2.2.2(vi)]**

$$\cos(\phi) = \frac{\tanh(\omega'_k)}{\tanh(R'/2)}, \quad \cos(\phi + \theta) = \tanh(\omega'_k).$$

Out of these relations we get, by elementary transformations,

$$\sin(\theta) = \frac{\tanh(\omega'_k)}{\tanh(R'/2)} \cdot \frac{1}{\cosh(\omega'_k)} \left\{ 1 - \sqrt{1 - \frac{\cosh(\omega'_k)^2}{\cosh(R'/2)^2}} \right\}.$$

Using that for $0 \leq x \leq 1$ we have $1 - \sqrt{1 - x} \leq x$ and that $\theta < \pi/2 \sin(\theta)$ we further obtain

$$\theta < \frac{\pi \sinh(\omega'_k)}{\sinh(R')}.$$

We now introduce additional white points on the arc that goes from C to $D \in \eta$ on the circle with radius R' and center \tilde{q} as shown in Figure 12. We position them in such a way that the successive arcs between them have lengths ε . Let z_2 be the number of these points, not counting C . Then $z_2 \leq \frac{\theta}{\varepsilon} \sinh(R') < \frac{\pi}{\varepsilon} \sinh(\omega'_k)$. Using that $\cosh(\omega'_k) = \tanh(\varepsilon/2) \coth(\sigma_k/2) \leq \frac{\varepsilon}{\sigma_k}$, **formula [3, 2.3.1(iv)]** applied to the trirectangle with sides $\varepsilon/2, \omega'_k, \sigma_k/2, \omega_k$ we further get

$$z_2 < \frac{\pi}{\sigma_k}.$$

Now $z_1 + z_2$ is the number of white points to the right of B (Figure 12) and there is the same number of similar points on the left. Drawing the connecting geodesic arcs from \tilde{q} to these points (including the arc from \tilde{q} to B) and projecting them from the universal covering to S we get the *terminal segments at q* .

So far we have assumed that $R \geq 2\omega'_k$. If $R < 2\omega'_k$, then $z_1 = 0$ and the number of white points on the circular arc of radius R is smaller than or equal to the number of such points we would get on the arc of radius $2\omega'_k$. For the latter we have already found the bound $\frac{\pi}{\sigma_k}$. Hence again $z_2 < \frac{\pi}{\sigma_k}$.

Summing up we have the following.

Lemma 4.5. *At any white point q on the boundary of \mathcal{C}_k'' there are at most $\frac{2}{\sigma_k}(R + 4)$ terminal segments.*

Finally, if q is a white point on the boundary of a reduced cusp neighborhood $\mathcal{P}_i^!$ then we need only one terminal segment, namely the geodesic arc of length R at q that is orthogonal to the horocycles.

4.4. Constructing the model strands

We begin by assigning to each collar \mathcal{C}_k'' a *winding number* τ_k with respect to \bar{c} as follows. If \mathcal{C}_k'' contains no traversing arc we set $\tau_k = 0$. Otherwise we let \mathcal{T}_k be the longest traversing arc in \mathcal{C}_k'' , project it orthogonally to a parametrized curve $\hat{\mathcal{T}}_k$ on β_k (that may go around β_k many times) and set

$$\tau_k = s_k \left\lceil \frac{\ell(\hat{\mathcal{T}}_k)}{\ell(\beta_k)} \right\rceil, \quad (21)$$

where $[x]$ for $x \in \mathbb{R}$ denotes the largest integer $\leq x$ and $s_k \in \{-1, 1\}$ is the orientation, i.e., $s_k = 1$ if $\hat{\mathcal{T}}_k$ winds around β_k in the positive sense (with respect to a fixed orientation of β_k) and $s_k = -1$ otherwise. At some later point we shall simultaneously unwind all traversing arcs in \mathcal{C}_k'' by applying a Dehn twist D_k of order τ_k along β_k in the “unwinding direction”: if \mathcal{T} traverses \mathcal{C}_k'' from, say A to A' , then $D_k(\mathcal{T})$ is the geodesic arc from A to A' in \mathcal{C}_k'' that is homotopic (with fixed endpoints) to the curve \mathcal{T}' that goes along \mathcal{T} from A to β_k , then $|\tau_k|$ times around β_k in the opposite direction of $\hat{\mathcal{T}}_k$ and after that along \mathcal{T} to A' . It follows from this construction that the orthogonal projection of $D_k(\mathcal{T}_k)$ on β_k has length $< \ell(\beta_k)$. Since the traversing arcs are pairwise disjoint we conclude that all of them have this property, i.e., *any $D_k(\mathcal{T})$ winds less than once around \mathcal{C}_k'' .*

The model strand for \bar{c} is constructed in a similar fashion as in Section 3, though with a modification for the traversing arcs and the terminal arcs. We proceed in three steps. First we split \bar{c} into a product

$$\bar{c} = c_0 c_1 c_2 c_3 \cdots c_{2J+1} c_{2J+2}$$

($J \geq 0$), where c_0 and c_{2J+2} are either terminal arcs (Section 4.3) or point curves depending on whether or not \bar{c} begins, respectively ends with a terminal arc; the parts $c_1, c_3, \dots, c_{2J+1}$ are outside the dotted collars and dotted cusp neighborhoods; the parts c_{2j} , for $j = 1, \dots, J$, comprise all the traversing arcs, each of them traversing some dotted collar \mathcal{C}_{k_j}'' . Here we are using that a complete simple geodesic cannot enter and leave a collar on the same side (Lemma 2.3) and that, similarly, it cannot enter and leave a cusp neighborhood. To deal with the windings separately we first “unwind” \bar{c} and also leave out c_0 and c_{2J+2} setting

$$\check{c} = c_1 \check{c}_2 c_3 \check{c}_4 c_5 \cdots c_{2J-1} \check{c}_{2J} c_{2J+1},$$

where $\check{c}_{2j} = D_{k_j}(c_{2j})$, $j = 1, \dots, J$. Now \check{c} traverses a succession of Voronoi cells and we associate to it a polygonal curve $P(\check{c})$ by the same procedure as in Section 3.3 with a minor adaptation: if c_0 is a terminal arc, say $c_0 \subset \mathcal{C}_\ell''$, then $P(\check{c})$ has its initial point on the boundary of \mathcal{C}_ℓ'' (as though \mathcal{C}_ℓ'' was the 0-th Voronoi cell that \check{c} goes along). This initial point is then one of the white points. Similarly, if c_{2J+2} is a terminal arc in some \mathcal{C}_k'' , then the endpoint of $P(\check{c})$ is a white point on the boundary of \mathcal{C}_k'' .

The curves \check{c} and $P(\check{c})$ have nearby endpoints: there is a geodesic arc \check{u} of length $\leq 2\varepsilon$ leading from the initial point of \check{c} to the initial point of $P(\check{c})$ and a geodesic arc \check{v} of length $\leq 2\varepsilon$ leading from the endpoint of $P(\check{c})$ to the endpoint of \check{c} . The paths \check{c} and $\check{u}P(\check{c})\check{v}$ are homotopic.

In the next step we apply to $P(\check{c})$ the reversed Dehn twists D_k^{-1} , $k = 1, \dots, h$. The resulting curve $\widehat{P}(\check{c}) = D_1^{-1} \dots D_h^{-1}(P(\check{c}))$ has the same endpoints as $P(\check{c})$ and $\check{u}\widehat{P}(\check{c})\check{v}$ is homotopic to the curve $c_1c_2c_3c_4 \dots c_{2J+1}$.

In the final step we add the possible terminal segments. We describe this for c_{2J+2} , the procedure for c_0 being the same. If c_{2J+2} is a point curve we set $\check{c}_{2J+2} = c_{2J+2}$. Now assume that c_{2J+2} is a terminal arc in some dotted collar \mathcal{C}_k'' . Then the initial point of c_{2J+2} lies within distance $\leq \varepsilon/2$ of the endpoint, say q , of $\widehat{P}(\check{c})$ on the boundary of \mathcal{C}_k'' and we first extend c_{2J+2} to a longer geodesic arc c'_{2J+2} in such a way that if we lift it to the universal covering, then the other endpoint lies within distance $\leq \varepsilon/2$ of some white point, say \tilde{r} , on the arc BCD as depicted in Figure 12 (or its symmetric image across $\tilde{q}B$), where the dotted line is c'_{2J+2} . The geodesic from \tilde{q} to \tilde{r} is the lift of a terminal segment and we let \check{c}_{2J+2} be this terminal segment. There is a connecting arc v of length $\leq \varepsilon/2$ from the endpoint of \check{c}_{2J+2} to the endpoint of c'_{2J+2} , and the curves $(\check{v})^{-1}\check{c}_{2J+2}v$ and c'_{2J+2} are homotopic. Finally, if c_{2J+2} is a terminal arc in some dotted cusp neighborhood \mathcal{P}_i'' , then we proceed similarly except that we take the extension c'_{2J+2} to be of length R . In this latter case \check{c}_{2J+2} is a geodesic arc orthogonal to the horocycles.

We now let \bar{c}' be the extended geodesic arc $\bar{c}' = c'_0c_1c_2c_3c_4 \dots c_{2J+1}c'_{2J+2}$ (where c'_0 is defined in the same way as c'_{2J+2}) and define the model strand m^c to be the geodesic arc in the homotopy class of $\check{c}_0\widehat{P}(\check{c})\check{c}_{2J+2}$. Then there is a connecting arc u of length $\leq 2\varepsilon$ from the initial point of \bar{c}' to the initial point of m^c , and the already described connecting arc v of length $\leq 2\varepsilon$ from the endpoint of m^c to the endpoint of \bar{c}' . By the aforementioned homotopies the curves um^cv and \bar{c}' are homotopic. In the disk D we have therefore the same estimate for the distance between c and m_c as in (12) and so we get the following result for our model strands m^c and their components m_c in D :

Lemma 4.6. *Any component c of the Birman-Series set in D lies in a tubular neighborhood*

of radius $w_R := 3e^{-R}$ of the model strand m_c .

4.5. Estimating the number of model strands

Lemma 4.7. *Let \check{L} be the number of path edges of $P(\check{c})$ and τ_1, \dots, τ_h the winding numbers of \bar{c} at β_1, \dots, β_h as in (21), then*

$$\frac{\varepsilon}{24} \check{L} + \sum_{k=1}^h \ell(\beta_k) |\tau_k| \leq \ell(\bar{c}) + 3\varepsilon.$$

Proof. For $j = 0, \dots, J$, any Voronoi cell (including the special ones) that is crossed by c_{2j+1} has its center at distance $\leq \varepsilon$ from c_{2j+1} (by the completeness of the ε net in S'). By the same area argument as in the proof of Lemma 3.6 it follows that c_{2j+1} crosses at most $\frac{4}{\varepsilon}(\ell(c_{2j+1}) + 3\varepsilon)$ times some Voronoi cell. Each \check{c}_{2j} , for $j = 1, \dots, J$, crosses an additional number of at most 4 special Voronoi cells. This is so because \check{c}_{2j} winds at most once around the corresponding collar and its initial and final cells are on the accounts of c_{2j-1} and c_{2j+1} , respectively. Altogether \check{c} crosses at most $\frac{4}{\varepsilon} \sum_{j=0}^J (\ell(c_{2j+1}) + 3\varepsilon) + 4J$ times some Voronoi cell and this multiplied by 6 is an upper bound of \check{L} :

$$\check{L} \leq \frac{24}{\varepsilon} \sum_{j=0}^J (\ell(c_{2j+1}) + 3\varepsilon) + 24J. \quad (22)$$

For any $k = 1, \dots, h$, such that $\tau_k \neq 0$ there is the longest traversing arc \mathcal{T}_k in \mathcal{C}_k'' which is one of the c_{2j} say $\mathcal{T}_k = c_{2j(k)}$. By (20) its orthogonal projection $\hat{c}_{2j(k)}$ on β_k has length satisfying $\ell(\beta_k) |\tau_k| \leq \ell(\hat{c}_{2j(k)}) < \ell(c_{2j(k)}) - 4/3 \leq \ell(c_{2j(k)}) - 4\varepsilon$. Using that any other traversing arc has length $\ell(c_{2j}) > 4\varepsilon$ we get

$$\sum_{k=1}^h \ell(\beta_k) |\tau_k| \leq \sum_{j=1}^J (\ell(c_{2j}) - 4\varepsilon) \quad (23)$$

With (22) and (23) the proof is complete. \square

In the next lemma the constant G stems from Corollary 4.4. In the case $n = 0$ it is the same as in (8), Section 3.6.

Lemma 4.8. *Set $G = \frac{97}{\varepsilon^2}(g - 1 + \frac{n}{2})$ and*

$$\mathcal{N}(R) = \frac{1}{10} \frac{m^G}{G^G} R^G \frac{1}{\ell(\beta_1) \cdots \ell(\beta_h)},$$

with $m = \frac{134}{\varepsilon}$. Then for given $R > G$ there are at most $\mathcal{N}(R)$ model strands in D .

Proof. We first count how many model strands arise from extensions \bar{c} that have both endpoints outside the dotted collars. We shall say that these curves belong to the *first category*. For any of them the model strand m^c is uniquely determined by the following three sets of data: the winding numbers τ_1, \dots, τ_h , the numbers n_1, \dots, n_V of path edges of $P(\check{c})$ on each of the V edges of the cell decomposition \mathcal{V} (which we assume enumerated from 1 to V in some way) and the selection of the initial and end path edge. By Lemma 4.7 we have, noting that $\check{L} = n_1 + \dots + n_V$, and $\ell(\bar{c}) \leq 2R - 4\varepsilon$,

$$\frac{\varepsilon}{24}(n_1 + \dots + n_V) + \sum_{k=1}^h \ell(\beta_k) |\tau_k| \leq 2R. \quad (24)$$

The number of model strands in the present case is thus bounded above by $4\check{L}^2 2^h$ times the number of strings $(n_1, \dots, n_V, |\tau_1|, \dots, |\tau_h|)$ satisfying inequality (24). By Corollary B.2 this number is bounded above by

$$B := \frac{1}{(V+h)! \left(\frac{\varepsilon}{24}\right)^V \ell(\beta_1) \dots \ell(\beta_h)} (2R + \gamma)^{V+h}, \quad (25)$$

where $\gamma = \frac{\varepsilon}{24}V + \ell(\beta_1) + \dots + \ell(\beta_h)$. Hence, the upper bound $4\check{L}^2 2^h B$ for the number of model strands with endpoints outside the dotted collars.

We also note using (17) and Corollary 4.4 that $h \leq 3(g-1 + \frac{n}{2})$ and hence,

$$V+h \leq G-5, \quad \gamma < \frac{1}{50}R. \quad (26)$$

We extend our first category of curves by allowing, in addition, the endpoints to lie in dotted collars that contain traversing arcs. For this extended category we have the previous bound multiplied by the number of possible choices of the terminal segments that are attached to $\widehat{P}(\check{c})$. Any such segment must be homotopic with fixed endpoints to a curve that does not intersect $\widehat{P}(\check{c})$ except at the endpoints. One of the latter is the attachment to $\widehat{P}(\check{c})$, the other is among the four white points on the opposite boundary component of the corresponding dotted collar. Hence, at either end of $\widehat{P}(\check{c})$ there are at most 5 possible choices for the terminal segment. This yields the bound $100\check{L}^2 2^h B$ for the number of curves in the extended first category.

The *second category* consists of the cases where exactly one endpoint of \bar{c} lies in a dotted collar that contains *no* traversing arcs or where both endpoints lie in the same dotted collar that contains no traversing arcs. We shall count the arising model strands for the cases for which this collar is \mathcal{C}_k'' and then take the sum for $k = 1, \dots, h$. Now for \bar{c} with initial point in \mathcal{C}_k'' (and possibly the end point also but without traversing arcs in \mathcal{C}_k'') the attachment

point for the terminal segment that is glued to the beginning of $\widehat{P}(\check{c})$ is uniquely determined by the sequence (n_1, \dots, n_V) , but the previous bound of 5 for the possible directions is now replaced by $\frac{8}{\ell(\beta_k)}(R+4)$ (Lemma 4.5). At the same time, since $\tau_k = 0$, inequality (24) is now replaced by $\frac{\varepsilon}{24}(n_1 + \dots + n_V) + \sum_{j=1, j \neq k}^h \ell(\beta_j) |\tau_j| \leq 2R$, and Corollary B.2 yields the bound

$$B' := \frac{\ell(\beta_k)}{(V+h-1)! \left(\frac{\varepsilon}{24}\right)^V \ell(\beta_1) \cdots \ell(\beta_h)} (2R + \gamma)^{V+h-1}.$$

for the number of strings that satisfy it. If the second endpoint of \bar{c} happens to lie in \mathcal{C}_k'' also, then both attachment points are determined by (n_1, \dots, n_V) but for given direction of the terminal segment at the beginning there are at most 5 directions for the terminal segment at the end, owing to the fact that both terminal arcs of \bar{c} in \mathcal{C}_k'' lie on the same simple geodesic. It follows that for \mathcal{C}_k'' there are at most $10\check{L}2^{h-1}B'\frac{8}{\ell(\beta_k)}(R+4) \leq 40\check{L}2^h(V+h)B$ possible cases, and summing up for $k = 1, \dots, h$ we get the upper bound $40h\check{L}2^h(V+h)B$ for the number of model strands arising from curves in category 2.

The *third* and final category consists of the cases where the two endpoints of \bar{c} lie in distinct dotted collars that both contain no traversing arcs. A similar argument as before shows that there are at most $8h^22^h(V+h)(V+h-1)B$ model strands arising from this last category.

Let now N be the sum of the bounds for the three categories. By Lemma 4.7 and since $\ell(\bar{c}) \leq 2R - 4$ (Section 4.3) we have $\check{L} \leq \frac{48}{\varepsilon}R$ (26). Furthermore, by the hypothesis of the lemma, $G < R$ and therefore $V+h < R$. Hence, allowing rough estimates at this point,

$$N \leq 100 \left(\frac{48}{\varepsilon}\right)^2 R^2 \left\{1 + \frac{1}{10}h + \frac{1}{100}h^2\right\} 2^h B \leq 100 \left(\frac{48}{\varepsilon}\right)^2 R^2 3^h B.$$

Applying to B that $\gamma < \frac{1}{50}R$ (26) and using that $x! \geq x^x e^{-x}$ for $x > 0$ we get

$$B \leq \frac{1}{\ell(\beta_1) \cdots \ell(\beta_h)} \left(\frac{24}{\varepsilon}\right)^V \left(2 + \frac{1}{50}\right)^{V+h} \left(\frac{eR}{V+h}\right)^{V+h}.$$

Since the function $x \rightarrow (eR/x)^x$ is monotone increasing for $x \in [1, R]$ and, by (26), $V+h \leq G-2 < R$, the last factor has the bound $(eR/(V+h))^{V+h} \leq (eR/(G-2))^{G-2}$. The bound in the lemma now follows by elementary simplification using that $G > 400$. \square

4.6. Finalizing the proof

Theorem 4.1 is now proved by the same argument as in the proof of Theorem 3.1: the given disc B_ρ has area $\geq \pi\rho^2$; for any model strand m_c in B_ρ the part $m_c^{2w_R}$ of the tubular neighborhood of radius $2w_R$ that lies in the disc has area $< 9\rho w_R$ (see (13)),

where $w_R = 3e^{-R}$. By Lemma 4.8 there are at most $\mathcal{N}(R)$ such strands and by Lemma 4.6 the tubular neighborhoods of radius w_R around them contain the Birman-Series set $\cap B_\rho$. We thus get a point p in B_ρ at distance $\geq w_R$ from it if we set R to be a solution $\geq G$ to the equation

$$\frac{1}{\rho\sigma} \frac{m^G}{G^G} R^G = e^R.$$

This is (14) with ρ replaced by $\rho\sigma$. By Lemma 3.10 this solution satisfies

$$R \leq 2 \log\left(\frac{1}{\rho\sigma}\right) + 2G \log(m) \quad (27)$$

and, analogously to (16), we get the bound $w_R \geq 3\rho^2\sigma^2 e^{-2G \log m}$ with G and m as in Lemma 4.8. This completes the proof. \square

5. SYSTOLE INDEPENDENT BOUNDS

From a geometric point of view collars around very small geodesics are similar to pairs of cusps. This suggests that there should also exist a version of Theorem 3.1 with bounds that are independent of the systole. Here we show that this is indeed possible. However, the constants arising from our approach become extremely small.

The main idea is that if the width of a collar is sufficiently large with respect to R then it cannot contain traversing arcs and, furthermore, the number of possible directions of terminal segments (c.f. Lemma 4.5) at any white point on the boundary is just equal to 1 as in the case of a cusp.

In Figure 13 we calculate how much larger than R the width must be. The figure depicts again part of the universal covering of the surface S and is in correspondence with the earlier Figure 12. The distance from the white point \tilde{q} to the lift $\tilde{\beta}_k$ of the small geodesic β_k is equal to $R + H$ with H to be determined and $\sigma_k = 1/4\ell(\beta_k)$. The geodesic ray η issued at \tilde{q} is asymptotic to $\tilde{\beta}_k$ and forms an angle θ with the vertical geodesic from \tilde{q} orthogonally to $\tilde{\beta}_k$. We now determine H in such a way that η contains the hypotenuse of a geodesic triangle $\tilde{q}CD$ with right angle at C and small sides R and ε . The latter is almost identical with the arc CD from C to η on the circle of radius R centered at \tilde{q} . Speaking in terms of Figure 12 we have $B = C$ and there is only one white point on the curve BCD (and its symmetric image across $\tilde{q}C$) and thus only one choice for the direction of a terminal segment at q in S . The necessary distance H is determined by the following triangle formulas (**formula [3, 2.2.2(iv)]**)

$$\sinh(R) = \tanh(\varepsilon) \cot(\theta), \quad \sinh(R + H) = \cot(\theta). \quad (28)$$

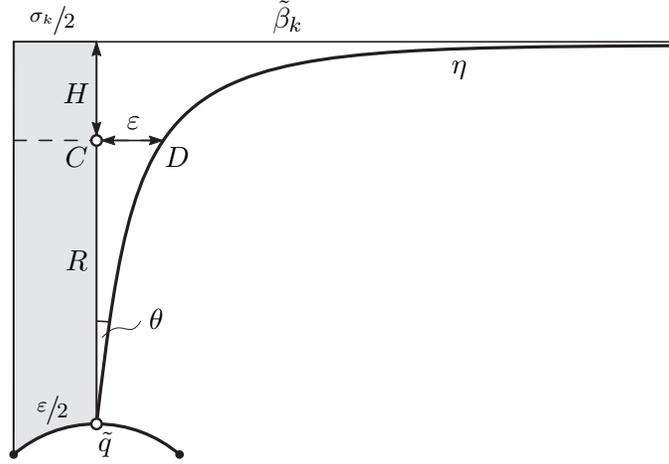


Figure 13: Asymptotic geodesic depicted in the universal covering

By the formula for the grey shaded trirectangle already used prior to (20) (with $\omega'_k = R+H$) we have $\cosh(R+H) = \tanh(\varepsilon/2) \coth(\sigma_k/2)$. For H determined by (28) the length of β_k , with negligible error, is equal to $16 \tanh(\varepsilon/2) \tanh(\varepsilon) e^{-R}$. Thus, if we define

$$\mathcal{L}(R) = 4\varepsilon^2 e^{-R} \quad (29)$$

then for given R any collar \mathcal{C}_k'' with $\ell(\beta_k) \leq \mathcal{L}(R)$ may be dealt with as though it was a pair of cusps.

We now start an iteration beginning by setting

$$\ell_0 = 1, \quad R_0 = 2 \log\left(\frac{1}{\rho}\right) + 2G \log(m),$$

where ρ , G , m are as in Theorem 4.1 and Lemma 4.8. The value for R_0 stems from (27) and is the bound for R in the proof of Theorem 4.1 that holds if $h = 0$, i.e., in the case where S has no small geodesics. By what we have said above, this bound is also valid if all small geodesics on S have lengths $\leq \mathcal{L}(R_0)$, and so under this weaker hypothesis we still have the lower bound $3e^{-R_0}$ for the largest distance to the Birman-Series set in the disc B_ρ .

In the first iteration step we set $\ell_1 = \mathcal{L}(R_0)$ and set R_1 equal to the right hand side of (27) for the special case $\sigma = \ell_1$. This is then the bound for R in the proof of Theorem 4.1 that holds if $h = 1$, and the same bound is valid in the more general case where all small geodesics different from β_1 have lengths $\leq \mathcal{L}(R_1)$. In this way we continue getting two sequences sequence $\ell_0, \ell_1, \ell_2, \dots$ and R_0, R_1, R_2, \dots , with the iteration scheme

$$\ell_k = \mathcal{L}(R_{k-1}) = 4\varepsilon^2 e^{-R_{k-1}}$$

$$R_k = 2 \log \frac{1}{\rho} + 2 \log \frac{1}{\ell_1} + \cdots + 2 \log \frac{1}{\ell_k} + 2G \log(m)$$

and at each step the result that some point in B_ρ has distance $\geq w_{R_k} = 3e^{-R_k}$ to the Birman-Series set provided that on S all small geodesics different from β_1, \dots, β_k are shorter than $\mathcal{L}(R_k)$. Now the recursion for R_k is equivalent to

$$R_k = 3R_{k-1} + 2 \log \frac{1}{4\varepsilon^2}$$

and R_k has the closed form

$$R_k = 3^k R_0 + (3^k - 1) \log \frac{1}{4\varepsilon^2}.$$

But the iteration stops at $k = 3g - 3 + n$ at the latest because this is the maximal possible number of small geodesics for S of genus g with n cusps. Hence we have the following.

Theorem 5.1. *Let S be a hyperbolic surface of genus g with n cusps and B_ρ a disk of radius ρ in the ε -thick part of S , where $0 < \rho < \varepsilon \leq 1/3$. Then there exists a point $p \in B_\rho$ such that*

$$\text{dist}(p, \mathcal{BS}) \geq e^{-3^\kappa R},$$

where $\kappa = 3g - 3 + n$ and $R = 2 \log \frac{1}{\rho} + M(g - 1 + n/2)$ with $M = \frac{195}{\varepsilon^2} \log \frac{134}{\varepsilon}$.

APPENDIX A: Hyperbolic triangles

For the convenience of the reader we gather a number of properties of hyperbolic triangles inscribed in a circle that are certainly well known but not easily accesible in the literature.

Lemma A.1. *Consider, for given $\varepsilon > 0$, a hyperbolic triangle with sides of lengths $\geq \varepsilon$ inscribed in a circle of radius $\leq \varepsilon$. Then all angles are bounded from below by φ_ε , where*

$$\cot \frac{\varphi_\varepsilon}{2} = \cosh(\varepsilon) \cdot \left\{ \sqrt{1 + 2 \cosh \varepsilon} + \sqrt{2 + 2 \cosh \varepsilon} \right\}. \quad (30)$$

For the proof we shall show that the smallest possible angle is φ_ε and is achieved in the case as shown on the left hand side in Figure 14, where we have an isosceles triangle ABC with base AB of length ε inscribed in a circle of center O and radius ε . An intermediate step in the proof is the following property that does not hold in Euclidean geometry.

Lemma A.2. *Let ABC be an arbitrary hyperbolic triangle inscribed in some circle of radius $\rho > 0$. If we allow C to move on the circle without crossing A or B , then the angle γ at C becomes minimal if and only if $\text{dist}(A, C) = \text{dist}(B, C)$.*

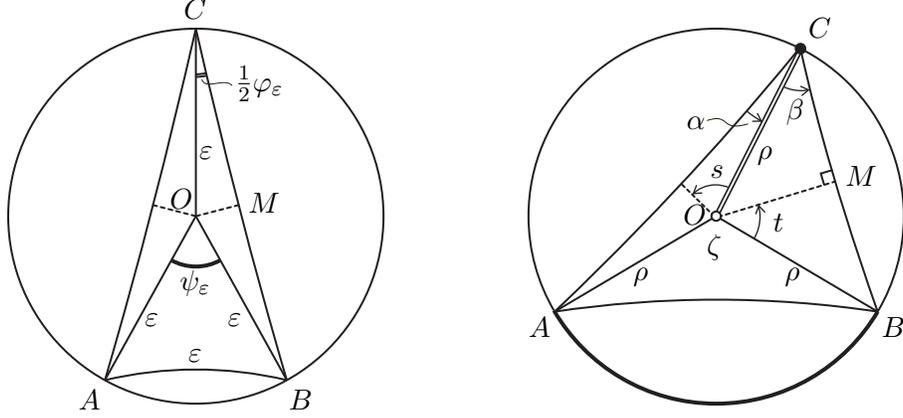


Figure 14: Two different triangles

Proof. We let $\zeta \in]0, 2\pi[$ be the measure of the angular region between OA and OB into which C is not allowed to enter. As a parameter for the position of C we then take the angle $2t \in]0, 2\pi - \zeta[$ from OB to OC . The orthogonal line from O to BC decomposes the triangle OBC into two isometric right-angled triangles OMC , OMB with the oriented angle $\beta = \beta(t)$ at C , the orientation being from CO to CB .

In a similar way we let $2s \in]0, 2\pi - \zeta[$ be the angle between OC and OA . The orthogonal line from O to AC decomposes the triangle OAC into two isometric right-angled triangles with oriented angle $\alpha = \alpha(s)$ at C , the orientation being from CA to CO . With these orientation conventions we have

$$2t + 2s + \zeta = 2\pi$$

For the right-angled triangles the following formulas hold:

$$\cosh \rho = \cot \beta \cot(t); \quad \cosh \rho = \cot \alpha \cot(s).$$

Indeed, for the configuration as in Figure 14, where $s, t, \alpha, \beta \in]0, \frac{\pi}{2}[$ these are instances of **formula [3, 2.2.2(ii)]**, and using that $\cot(\pi - s) = -\cot(s)$ we easily see that the formulas remain valid in the cases where $\alpha \leq 0$ or $\beta \leq 0$. One also may check that in all cases the angle γ of triangle ABC at C satisfies

$$\gamma = \alpha + \beta.$$

A straightforward computation involving trigonometric identities (including $\tan(s) + \tan(t) = \frac{\sin(s+t)}{\cos(s)\cos(t)}$) now yields the following formula for γ as a function of $t \in]0, \pi - \frac{\zeta}{2}[$,

$$\cot \gamma(t) = \cot \frac{\zeta}{2} \cdot \cosh \rho - \frac{1}{2} \frac{\sinh^2 \rho}{\sin \frac{\zeta}{2} \cosh \rho} \left\{ \cos(2t + \frac{\zeta}{2}) + \cos \frac{\zeta}{2} \right\}.$$

Thus, γ becomes minimal when $2t = \pi - \frac{\zeta}{2}$. □

Proof of Lemma A.1. By Lemma A.2 it suffices to look at triangles ABC that are isosceles at C . If we move A, B on the given circle towards each other until their distance is ε the angle γ decreases. If we then increase the height of the isosceles triangle keeping the base AB fixed until the circumcircle reaches radius ε the angle γ decreases again. Hence, the comparison triangle with the minimal angle φ_ε as on the left hand side in Figure 14.

To compute φ_ε we begin with the angle ψ_ε at O of the equilateral triangle OAB . By formula [3, 2.2.2(iii)] we have $\sinh \frac{\varepsilon}{2} = \sin \frac{\psi_\varepsilon}{2} \cdot \sinh \varepsilon$ or equivalently, using the half angle formulas for cot and sinh,

$$\cot \frac{\psi_\varepsilon}{4} = \sqrt{1 + 2 \cosh \varepsilon} + \sqrt{2 + 2 \cosh \varepsilon}. \quad (31)$$

In the right-angled triangle OMC the angles at O and C are $\frac{1}{2}\pi - \frac{1}{4}\psi_\varepsilon$ and $\frac{1}{2}\varphi_\varepsilon$, respectively. By the formula used earlier,

$$\cosh \varepsilon = \cot \frac{\varphi_\varepsilon}{2} \cdot \cot\left(\frac{\pi}{2} - \frac{\psi_\varepsilon}{4}\right) = \cot \frac{\varphi_\varepsilon}{2} \cdot \tan \frac{\psi_\varepsilon}{4},$$

which implies the formula of the lemma. \square

For completeness we add an upper bound on the angles of an inscribed triangle. The argument for this is easy: the largest angle is reached for the triangle ABD inscribed in a circle of radius ε with sides AB and BD of length ε . Triangles OAB and OBD are equilateral with the interior angles ψ_ε and so we have

Lemma A.3. *Consider, for given $\varepsilon > 0$, a hyperbolic triangle with sides of lengths $\geq \varepsilon$ inscribed in a circle of radius $\leq \varepsilon$. Then all angles are bounded from above by $2\psi_\varepsilon$, where ψ_ε is as in (31).*

APPENDIX B: A combinatorial lemma

To distribute up to L identical objects into K distinct boxes is possible in $\binom{L+K}{L}$ different ways; a simple bound is $\frac{(L+K)^K}{K!}$. Here we prove a bound for the case for distributions in packets.

Lemma B.1. *For any string of positive integers $\vec{s} = (s_1, \dots, s_K) \in (\mathbb{N} \setminus \{0\})^K$ and any $L \in \mathbb{N} \setminus \{0\}$ we denote by $\chi_{\vec{s}}(L)$ the number of ordered sequences $(n_1, \dots, n_K) \in \mathbb{N}^K$ satisfying*

$$n_1 s_1 + \dots + n_K s_K \leq L.$$

Then

$$\chi_{\vec{s}} \leq \frac{1}{K! s_1 \cdots s_K} (L + s_1 + \dots + s_K)^K.$$

Proof. By induction over K . For $K = 1$ the inequality is clear. For the step from K to $K + 1$ we abbreviate $s_{K+1} = s$, $n_{K+1} = n$. The possible values for n are $n = 0, 1, 2, \dots, \lfloor \frac{L}{s} \rfloor$ (largest integer $\leq \frac{L}{s}$). Now we observe that

$$\chi_{(s_1, \dots, s_{K+1})} = \sum_{n=0}^{\lfloor \frac{L}{s} \rfloor} \chi_{\vec{s}}(L - n s).$$

By the induction hypothesis, for any $n = 0, \dots, \lfloor \frac{L}{s} \rfloor$ the number of sequences n_1, \dots, n_K has the upper bound

$$\chi_{\vec{s}}(L - n s) \leq \frac{1}{K! s_1 \cdots s_K} (L - n s + A)^K,$$

where we have abbreviated $A = s_1 + \cdots + s_K$. It remains to prove that

$$\sum_{n=0}^{\lfloor \frac{L}{s} \rfloor} \frac{1}{K! s_1 \cdots s_K} (L - n s + A)^K \leq \frac{1}{(K+1)! s_1 \cdots s_K s} (L + A + s)^{K+1}.$$

Now

$$\begin{aligned} \sum_{n=0}^{\lfloor \frac{L}{s} \rfloor} (L + A - n s)^K &= \sum_{n=0}^{\lfloor \frac{L}{s} \rfloor} \int_{n-1}^n (L + A - n s)^K dt \leq \sum_{n=0}^{\lfloor \frac{L}{s} \rfloor} \int_{n-1}^n (L + A - t s)^K dt \\ &\leq \int_{-1}^{L/s} (L + A - t s)^K dt = \frac{1}{s(K+1)} ((L + A + s)^{K+1} - A^{K+1}) \\ &\leq \frac{1}{s(K+1)} (L + A + s)^{K+1} \end{aligned}$$

and the above inequality follows. \square

There is also a real valued version of the lemma:

Corollary B.2. *For any string of positive real numbers $\vec{\rho} = (\rho_1, \dots, \rho_K) \in \mathbb{R}^K$ and any $\lambda > 0$ we denote by $\chi_{\vec{\rho}}(\lambda)$ the number of ordered sequences $(n_1, \dots, n_K) \in \mathbb{N}^K$ satisfying*

$$n_1 \rho_1 + \cdots + n_K \rho_K \leq \lambda.$$

Then

$$\chi_{\vec{\rho}} \leq \frac{1}{K! \rho_1 \cdots \rho_K} (\lambda + \rho_1 + \cdots + \rho_K)^K.$$

Proof. Take $t > 0$ smaller than ρ_1, \dots, ρ_K and set $L = \lfloor \frac{\lambda}{t} \rfloor + 1$, $s_k = \lfloor \frac{\rho_k}{t} \rfloor$, for $k = 1, \dots, K$. Any string satisfying $n_1 \rho_1 + \cdots + n_K \rho_K \leq \lambda$ then also satisfies $n_1 s_1 + \cdots + n_K s_K \leq L$. We may thus apply Lemma B.1 and the corollary is obtained in the limit as $t \rightarrow 0$. \square

REFERENCES

- [1] Lipman Bers. An inequality for Riemann surfaces. In *Differential geometry and complex analysis*, pages 87–93. Springer, Berlin, 1985.
- [2] Joan S. Birman and Caroline Series. Geodesics with bounded intersection number on surfaces are sparsely distributed. *Topology*, 24(2):217–225, 1985.
- [3] Peter Buser. Geometry and spectra of compact Riemann surfaces, volume 106 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1992.
- [4] Peter Buser and Hugo Parlier, The distribution of simple closed geodesics on a Riemann surface, *Complex analysis and its applications, OCAMI Stud.*, 2. Osaka Munic. Univ. Press, Osaka, 2007.
- [5] Moira Chas, Non-abelian number theory and the structure of curves on surfaces, [arXiv:1608.02846](https://arxiv.org/abs/1608.02846), Preprint, 2016.
- [6] Viveka Erlandsson and Juan Souto, Counting curves in hyperbolic surfaces, *Geom. Funct. Anal.*, 26(3):729–777, 2016.
- [7] Heinz Huber. Zur analytischen Theorie hyperbolischer Raumformen und Bewegungsgruppen. *Math. Ann.*, 138: 1–26, 1959.
- [8] Troels Jørgensen. Simple geodesics on Riemann surfaces. *Proc. Amer. Math. Soc.*, 86(1):120–122, 1982.
- [9] Maryam Mirzakhani. Growth of the number of simple closed geodesics on hyperbolic surfaces, *Ann. of Math. (2)*, 168(1):97–125, 2008.
- [10] Maryam Mirzakhani, Counting mapping class group orbits on hyperbolic surfaces, [arXiv:1601.03342](https://arxiv.org/abs/1601.03342), Preprint, 2016.
- [11] Igor Rivin, Simple curves on surfaces, *Geom. Dedicata*, 87(1-3): 345–360, 2001.
- [12] Igor Rivin, Geodesics with one self-intersection, and other stories, *Adv. Math.*, 231(5): 2391–2412, 2012.
- [13] Jenya Sapir, Bounds on the number of non-simple closed geodesics on a surface., *Int. Math. Res. Not. IMRN*, 24:7499–7545, 2016.
- [14] Jenya Sapir, A Birman-Series type result for geodesics with infinitely many self-intersections, [arXiv:1609.00428](https://arxiv.org/abs/1609.00428), Preprint, 2016.

- [15] Peter Scott. Subgroups of surface groups are almost geometric. *J. London Math. Soc.* (2), 17(3):555–565, 1978.

Addresses:

Institute of Mathematics, EPFL, Switzerland

Email: peter.buser@epfl.ch

Department of Mathematics, University of Luxembourg, Luxembourg

Email: hugo.parlier@uni.lu