Groups Geom. Dyn. 17 (2023), 1–33
DOI 10.4171/GGD/699

A property of closed geodesics on hyperbolic surfaces

Max Neumann-Coto and Peter Scott

Abstract. We study closed geodesics on hyperbolic surfaces and give bounds for their angles of intersection and self-intersection, as well as for the size of the n-gons that they form, depending only on the lengths of the geodesics.

In this paper, we consider the geometry of closed geodesics on a hyperbolic surface M by looking at the geodesic lines that form their pre-images in $\mathbb{H}^2$. We are interested in finding lower bounds for the distances and the angles between these lines, as well as upper bounds for the size of the n-gons that they form, in terms of the lengths of the closed geodesics involved.

If l and m are two geodesic lines in the hyperbolic plane $\mathbb{H}^2$ that do not share a point at infinity, the orthogonal projection of l onto m has finite length, which depends on the distance between the lines if they are disjoint, or the angle of intersection if they meet. As there is a hyperbolic reflection interchanging l and m, the orthogonal projection of l onto m and the orthogonal projection of m onto l have the same length. The main result of the paper is the following.

Theorem 5.1. *Let γ be a closed geodesic on an orientable hyperbolic surface M, and let l and m be distinct geodesics in $\mathbb{H}^2$ above γ. Then the orthogonal projection of l onto m is shorter than the length of γ.*

The result in the case when l and m are disjoint is a consequence of the stable neighborhood theorem of Basmajian [1], of which we give a new proof. The case when l and m intersect is new and is a strong refinement of a result of Beardon [2, Theorem 11.6.8] (see also [3, Corollary IVb]), which we discuss later.

Theorem 5.1 gives bounds for the self-intersection angles of a hyperbolic geodesic γ that depend only on its length $l(\gamma)$ and not on the topology of the surface or the hyperbolic metric chosen. Recall that the *angle of parallelism* of a positive number a, denoted by $\Pi(a)$, is the third angle of a right hyperbolic triangle with one side of length a and two asymptotic parallel sides. In Corollary 5.2, we show that if γ is a closed oriented geodesic on an orientable hyperbolic surface M, the angle ϕ formed by two outgoing arcs of γ at any self-intersection point satisfies $\Pi(\frac{l(\gamma)}{2}) < \phi < \pi - \Pi(\frac{l(\gamma)}{4})$.

2020 Mathematics Subject Classification. Primary 51M10; Secondary 53C22, 57M60.
Keywords. Hyperbolic geodesic, projection, intersection angles.

These bounds hold for the angles formed by the axes of any two conjugate isometries of $\mathbb{H}^2$ that generate a free discrete subgroup of $\mathrm{SL}(2,\mathbb{R})$.

Theorem 5.1 also gives a bound for the size of the convex polygons in $\mathbb{H}^2$ formed by geodesic lines above a closed hyperbolic geodesic. In Corollary 5.3, we show that the triangles formed by the geodesic lines above γ in $\mathbb{H}^2$ have sides shorter than $l(\gamma)$, and the n-gons have sides shorter than $(n-2)l(\gamma)$.

The bound given in Theorem 5.1 is optimal in the sense that for each hyperbolic surface M, there is a sequence of geodesics in M and lines above them in $\mathbb{H}^2$ for which the ratio of the projection length to the geodesic length approaches 1. But this does not imply that the bound for the angles is optimal, and the bounds for the sides of n-gons might be far from optimal for $n > 3$.

Theorem 5.1 can be generalized to the case of two closed geodesics. In Corollary 5.4, we show that if γ and δ are closed geodesics on an orientable hyperbolic surface M, and if l and m are distinct geodesics in $\mathbb{H}^2$ above γ and δ, respectively, then the orthogonal projection of l onto m has length strictly less than $l(\gamma)+l(\delta)$. This bound is optimal too, in the sense that there are two sequences of closed geodesics γ_n and δ_n and lines l_n and m_n above them, for which the ratio between the length of the projection of γ_n to δ_n and the sum of the lengths of γ_n and δ_n approaches 1.

Corollary 5.4 gives lower bounds for the distance between two non-intersecting lines above γ and δ (which generalizes the stable neighborhood theorem of Basmajian to pairs of geodesics) and gives lower bounds for the intersection angles of intersecting lines. These bounds are not as good as the bounds given in [2] and [3], but we give them for completeness and because the proof of Corollary 5.4 in the case of intersecting geodesics is very simple (it uses only Lemma 2.1).

Let us now discuss the relations between the bounds given by Theorem 5.1 and previous results, in particular [2, Theorem 11.6.8] and [3, Corollary IVb]. Beardon and Gilman consider a pair of hyperbolic isometries A and B that generate a purely hyperbolic discrete subgroup of $\mathrm{SL}(2,\mathbb{R})$. If the translation lengths are $|A|$ and $|B|$ and the axes meet at an angle ϕ, they show that $\sin(\phi)\sinh(|A|/2)\sinh(|B|/2) > 1$. This lower bound on the angle ϕ is stronger than the bound from Corollary 5.7, but when A and B are conjugate, so that $|A| = |B|$, those results give the inequality

$$\sin(\phi) > \operatorname{csch}^2\left(\frac{|A|}{2}\right) = \frac{4}{(e^{\frac{|A|}{2}} - e^{-\frac{|A|}{2}})^2}.$$

On the other hand, in the case of a single closed geodesic, Theorem 5.1 implies that ϕ is larger than the parallelism angle of $\frac{|A|}{2}$. Hence

$$\sin(\phi) > \sin\Big(\Pi\Big(\frac{|A|}{2}\Big)\Big) = \operatorname{sech}\left(\frac{|A|}{2}\right) = \frac{2}{e^{\frac{|A|}{2}} + e^{-\frac{|A|}{2}}},$$

which gives a much stronger bound: for $|A| \geq 4\log(1+\sqrt{2})$ (a lower bound for the length a non-simple geodesic in any hyperbolic surface), the ratio of the two bounds is around 17, and as $|A|$ increases, the ratio grows exponentially to infinity.

Most of this paper is devoted to the proof of Theorem 5.1. We will use a mix of geometric and algebraic arguments to deal with the different cases, some of which are very simple and allow us to give some better bounds, but others are quite intricate and involve many steps. In Section 1, we prove some simple results about axes of elements of a group which acts on a tree. In Section 2, we consider the case of disjoint geodesics and some special cases of crossing geodesics. In Sections 3 and 4, we consider the cases when the surface is a three-punctured sphere or a once-punctured torus. Finally, in Section 5, we complete the proof of Theorem 5.1 and then deduce several consequences.

1. Groups acting on trees

Theorem 5.1 can be regarded as a geometric version of a simple result about the intersections of axes of elements of a group acting on a tree.

Let T denote the tree which is the universal cover of a finite graph X with a single vertex. Such a graph is called a *rose*. Thus the free group $G = \pi_1(X)$ acts freely on T with quotient X. After orienting each loop in the rose X, there is a naturally associated set S of free generators of G, one generator for each oriented loop. Thus, each oriented edge of T has an element of $S \cup S^{-1}$ associated to it, and any oriented edge path λ in T has a word w in $S \cup S^{-1}$ associated to it. Any element α of G can be expressed uniquely as a reduced word in $S \cup S^{-1}$. We are interested in the minimal length $L(\alpha)$ of all words representing conjugates of α. Such a word realizes this minimal length if and only if it is cyclically reduced. Any nontrivial element α of G has an axis A, which is an edge path in T that is preserved by α and on which α acts by a translation of length $L(\alpha)$. This is why we use the same letter L for this algebraically defined length as for the length of an edge path in T. The infinite reduced word associated to A, when suitably oriented, is made by concatenating copies of a cyclically reduced word w equal to a conjugate of α. We call this infinite word the unwrapping of w.

Lemma 1.1. *Let W be a reduced word in a free group G. If W contains two nontrivial subwords u and v and $u = v^{-1}$, then u and v cannot overlap.*

Proof. We have $u = s_1 s_2 \dots s_n$ for some s_i's in the generating set $S \cup S^{-1}$ of G. Thus $v = s_n^{-1} s_{n-1}^{-1} \dots s_1^{-1}$. Without loss of generality, we can assume that the overlap consists of an initial segment $s_1 \dots s_k$ of u and an ending segment $s_k^{-1} \dots s_1^{-1}$ of v. Denoting $s_1 \dots s_k$ by z, we see that z is a nontrivial reduced word that is equal to its own inverse, which is impossible. This completes the proof of the lemma. ■

In the next lemma, we give bounds for the length of the intersection of two axes in T with conjugate stabilizers.

Lemma 1.2. *Let G be a free group which acts freely on a tree T with quotient a rose, and naturally associated set S of free generators of G. Let α be a nontrivial element of G with* *axis A, and $L(\alpha) \geq 2$. Let α' and α'' be conjugates of α with axes A' and A'', respectively. Then the following inequalities hold:*

(1) *If A and A'' share an edge of T where the translation directions of α and α'' disagree, then $L(A \cap A'') < \frac{L(\alpha)-1}{2}$.*

(2) *If A and A' are distinct, then $L(A \cap A') < L(\alpha) - 1$.*

Proof. (1) Choose an orientation of A so that the associated infinite reduced word is the unwrapping $\widetilde{W}$ of a cyclically reduced word W equal to a conjugate of α. Let I denote the interval $A \cap A''$ regarded as a subinterval of A with the induced orientation, and let u denote the reduced word associated to I. Thus, $\widetilde{W}$ contains u as a subword, and also contains infinitely many translates of u by shifting by powers of W. Next, consider the axis A'' of α''. As α'' is conjugate to α, we know A'' is a translate of A, and we give A'' the induced orientation. Now the infinite reduced word associated to A'' is also equal to $\widetilde{W}$, and so also contains infinitely many copies of the subword u. If we cycle W to begin with u, we have $W = uv$ for some reduced word v, and $\widetilde{W} = \ldots uvuvuv \ldots$

As the translation directions of α along A and of α'' along A'' disagree, the infinite reduced word $\widetilde{W}$ must contain a copy of u^{-1}. Lemma 1.1 tells us that u and u^{-1} cannot overlap, so that u^{-1} must be a subword of the subword v of $\widetilde{W}$. Further, u^{-1} cannot contain the initial or the final letter of v, as this would contradict the fact that $\widetilde{W}$ is reduced. As u and u^{-1} have the same length, we conclude that $2L(u) \leq L(W) - 2$, so that $L(u) \leq \frac{L(w)-2}{2} = \frac{L(\alpha)-2}{2}$.

(2) Let I denote the interval $A \cap A'$. We will assume that $L(I) \geq 1$, as otherwise the result is trivial. Further, we can assume that the translation directions of α and α' agree on I, as otherwise the result follows from part (1).

If I has length at least $L(\alpha)$, and x denotes the initial point of I, then αx and $\alpha' x$ must be equal as each is a vertex of I with distance $L(\alpha)$ from x. Note that this uses our assumption that the translation directions of α and α' agree. It follows that α equals α', so that A equals A', which contradicts our hypothesis. Thus, we will suppose that I has length $L(\alpha) - 1$, and let x and y denote the endpoints of I. Now the translation lengths of α along A and of α' along A' are equal to $L(\alpha)$. Thus, by interchanging x and y if needed, we have that αx and $\alpha' x$ both have distance 1 from y. Thus $\alpha'^{-1}\alpha$ moves x distance 2. But $\alpha'^{-1}\alpha$ is a commutator as α' is a conjugate of α, and so either it is trivial or it moves any point distance at least 4. This is because a reduced word of length strictly less than 4 must be trivial or map to a non-zero element in the abelianization of G. We conclude that $\alpha'^{-1}\alpha$ is trivial, so that α and α' are equal and A is equal to A', which again contradicts our hypothesis. This contradiction shows that $L(I) < L(\alpha) - 1$, which completes the proof of part (2) of the lemma. ∎

The following examples show that the inequalities given in Lemma 1.2 are sharp. Let G be the free group of rank 2 generated by x and y, and let T be the standard 4-valent tree with vertices labelled by elements of G.

For part (1), we let $\alpha = xy^{-1}xy$ and $\alpha'' = x^{-1}\alpha x = y^{-1}xyx$, so that $L(\alpha) = 4$ and $\frac{L(\alpha)-1}{2} = \frac{3}{2}$. Thus $L(A \cap A'') < \frac{3}{2}$, and we claim that $L(A \cap A'') = 1$. As α and α'' are cyclically reduced, each of their axes, A and A'', passes through the vertex e of T. Now it is easy to check that A and A'' each contain the edge of T with vertices e and y^{-1}, and that indeed the translation directions of α along A and of α'' along A'' disagree on this edge. Hence $L(A \cap A'') = 1$, as required.

For part (2), we let $\alpha = xyx$, $\alpha' = x\alpha x^{-1} = xxy$, so that $L(\alpha) = 3$ and $L(\alpha) - 1 = 2$. Thus $L(A \cap A') < 2$, and we claim that $L(A \cap A') = 1$. As α and α' are cyclically reduced, each of their axes, A and A', passes through the vertex e of T. Now it is easy to check that A and A' each contain the edge of T with vertices e and x. Hence $L(A \cap A') = 1$, as required.

2. Hyperbolic geometry

Let M be an orientable hyperbolic surface and let γ be a closed geodesic on M. The universal cover $\widetilde{M}$ of M is isometric to the hyperbolic plane $\mathbb{H}^2$. If l and m are geodesic lines in $\mathbb{H}^2$ above γ, there is g in $\pi_1(M)$ such that $m = gl$. Let α denote a generator of the stabilizer of l, and consider the cover F of M with the fundamental group generated by α and g. As F is hyperbolic and $\pi_1(F)$ has two generators, F cannot be closed. As F is also orientable, it follows that it must be homeomorphic to a sphere with three disjoint discs removed, or a torus with a disc removed. Note that as g lies in $\pi_1(F)$, the geodesics l and m in $\mathbb{H}^2$ project to a single closed geodesic on F with the same length as γ. Hence, in order to prove Theorem 5.1, it suffices to handle the case when M is equal to F.

Geometrically, there are several distinct possibilities depending on whether each end of F is a cusp or contains a closed geodesic. For simplicity in most of our arguments below, we will consider only the special cases when the ends of F are cusps. It turns out that these are the most subtle cases. In Section 5, we will discuss how to prove Theorem 5.1 in general, using essentially the same arguments as in these special cases.

If F is a three-punctured sphere or a once-punctured torus, there is a pair of disjoint simple infinite geodesics which together cut F into an ideal quadrilateral Q. The three-punctured sphere is the double of an ideal triangle, and we choose two of the common edges of the triangles to be the geodesics which cut F into the ideal quadrilateral Q. The third common edge becomes one of the diagonals of Q. As the three-punctured sphere admits a reflection isometry which fixes the three common edges, it follows that Q admits a reflection isometry in this diagonal. Hence the diagonals of Q must meet at right angles. If F is a once-punctured torus, Q may not admit any reflection isometries.

Now we consider the universal cover $\widetilde{F}$ of F, which is isometric to $\mathbb{H}^2$ and is naturally tiled by copies of the quadrilateral Q, as in Figure 1. The tessellation obtained from a three-punctured sphere is very symmetric, but the tessellations obtained from a once-punctured torus need not be so symmetric. We will consider the tree T dual to these edges. The geodesics l and m are automatically transverse to these cutting edges and will

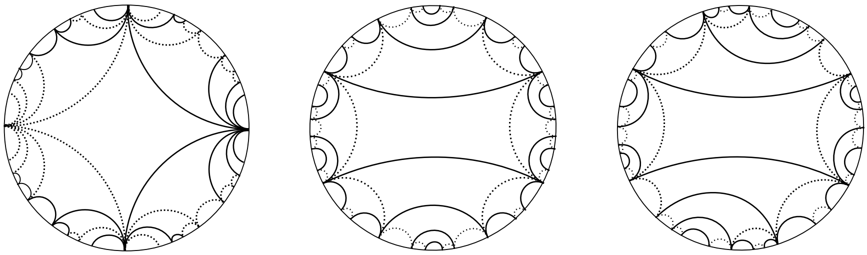


Figure 1. Tessellations of $\mathbb{H}^2$ by ideal quadrilaterals from a three-punctured sphere and from different once-punctured tori.

intersect each translate of Q in some (possibly empty) collection of embedded arcs. The group $\pi_1(F)$ acts freely on T with quotient that is a finite graph with a single vertex. There is a natural projection of $\widetilde{F}$ onto T which maps a (thin) collar neighborhood of each cutting geodesic onto an edge of T and maps the rest of each polygon to a vertex of T. The projections of the geodesics l and m to T are injective on the intersection with each collar neighborhood of a cutting arc. As two distinct geodesics in $\mathbb{H}^2$ cross at most once, it follows that the images of l and m traverse each edge of T at most once. Thus, the images in T of l and m are the axes of α and $g\alpha g^{-1}$, respectively.

Now we are in a position to apply our results from Section 1, but we will first consider some special cases of Theorem 5.1 which give some stronger bounds and can be settled using only geometric arguments.

We will say that the *bisector* of two disjoint geodesics λ and λ' in $\mathbb{H}^2$ is the unique geodesic Λ such that the reflection in Λ interchanges λ and λ'.

The first result is equivalent to the stable neighborhood theorem of Basmajian [1, Theorem 1].

Lemma 2.1. *Let γ be a closed geodesic on an orientable hyperbolic surface M, and let l and m be disjoint geodesics in $\mathbb{H}^2$ above γ. Then the orthogonal projection of the bisector of l and m onto m has length no greater than $l(\gamma)$.*

Proof. Let b be the bisector of l and m. We will show that the translates of b by the action of the stabilizer of l are disjoint. See Figure 2.

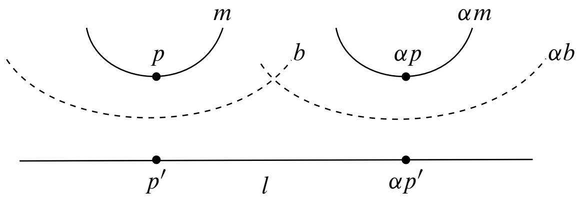


Figure 2. If l and m are disjoint geodesics.

If α is a generator of the stabilizer of l, and r is the reflection of $\mathbb{H}^2$ in b, then $r' = \alpha \circ r \circ \alpha^{-1}$ is the reflection of $\mathbb{H}^2$ in αb. If b and αb were not disjoint, then $r' \circ r$ would fix their intersection point, and so $r' \circ r$ would be an elliptic isometry. But this is impossible because $r' \circ r$ is an element of $\pi_1(M)$. To prove this, observe that $r' \circ r$ maps m to αm, preserving the orientations induced by γ. If p and p' are the closest points of the geodesics m and l, then r maps p to p', and r' maps $\alpha p'$ to αp. So $r' \circ r$ maps the point p in m to a point in αm at distance $l(\gamma)$ from αp. Therefore, $\alpha^{-1} \circ r' \circ r$ is a translation along m of length $l(\gamma)$, and so it lies in $\pi_1(M)$.

Let μ be the geodesic joining p and p', so μ crosses l, m and b orthogonally. Thus $\alpha\mu$ joins αp and $\alpha p'$ and crosses αl, αm and αb orthogonally. Now, if λ is the perpendicular bisector of the geodesic segment joining p' and $\alpha p'$, then reflection in λ preserves l and interchanges b and αb. It follows that λ is disjoint from b and αb, because if λ met b, it would have to meet αb in the same point, contradicting the fact that b and αb are disjoint. Hence λ crosses l orthogonally and separates b from αb. It follows that b lies between λ and $\alpha\lambda$, so the orthogonal projection of b onto l has length no greater than $l(\gamma)$, as required. ∎

Example 2.2. The bound given in the previous lemma is sharp. If M is a once-punctured torus for which the tessellation of $\mathbb{H}^2$ by quadrilaterals is symmetric, and γ is a longitude of M, then the bisector of two adjacent geodesics l and m above γ is an edge of a quadrilateral, which projects to an arc of l of length $l(\gamma)$.

At this point, we need a brief discussion of orientations of geodesics in $\mathbb{H}^2$. If two oriented geodesics in $\mathbb{H}^2$ are disjoint, it makes sense to say that they are coherently or oppositely oriented. Two oriented disjoint geodesics l and m are *coherently oriented* if for any geodesic λ which cuts both of them, they cross λ in the same direction, and they are *oppositely oriented* if for any geodesic λ which cuts both of them, they cross λ in opposite directions. Clearly, this does not depend on the geodesic λ. In particular, l and m are coherently oriented if and only if the orthogonal projection of l onto m is coherently oriented with m. (Note that the orthogonal projection of l onto m inherits a natural orientation from that of l, as it cannot consist of a single point unless l and m cross and do so orthogonally.) If we choose an orientation of a closed geodesic γ on a hyperbolic surface M, it induces an orientation of each geodesic in $\mathbb{H}^2$ above γ. If l and m are two disjoint such geodesics which are coherently oriented, they will remain coherently oriented if we reverse the orientation of γ, and the same applies if they are oppositely oriented. Thus we do not need to specify an orientation for γ in order to ask the question of whether l and m are coherently or oppositely oriented.

If l and m are crossing oriented geodesics in $\mathbb{H}^2$, the above definition of coherent orientation does not work, as the way they cross λ does depend on λ. But still, the orthogonal projection of l onto m is coherently oriented with m if and only if the orthogonal projection of m onto l is coherently oriented with l. (Unless l and m cross at right angles.)

Having discussed relative orientations for geodesics in $\mathbb{H}^2$, we also need to make clear the connection between this and the relative orientations of the corresponding axes in the tree T described above, which is dual to some family of cutting geodesics in $\mathbb{H}^2$. Let λ denote one of these cutting geodesics which meets disjoint geodesics l and m above a closed geodesic γ in a hyperbolic surface M. If e denotes the edge of T dual to λ, the axes in T to which l and m project must each contain e, and the directions in which l and m cross λ are the same as the direction in which the axes traverse e. Thus, if l and m are disjoint geodesics in $\mathbb{H}^2$ above γ and if their images A and B in T have a common edge, then l and m are coherently oriented if and only if A and B are coherently oriented on their intersection. However, if l and m are crossing geodesics above γ and if the axes A and B in T overlap, these axes may or may not be coherently oriented, and this may well depend on the choice of cutting geodesics in $\mathbb{H}^2$. In particular, it is possible that the orthogonal projection of l onto m is coherently oriented with m, while A and B overlap and have opposite orientations.

Having completed this discussion, we can give a special case of Theorem 5.1, for which we get an even lower bound on the orthogonal projection of l onto m. This is analogous to the result of part (1) of Lemma 1.2.

Lemma 2.3. *Let γ be a closed geodesic on an orientable hyperbolic surface M, and let l and m be disjoint geodesics in $\mathbb{H}^2$ above γ which are oppositely oriented. Then the orthogonal projection of l onto m has length strictly less than $l(\gamma)/2$.*

Remark 2.4. The fact that $\pi_1(M)$ is a discrete group of isometries of $\mathbb{H}^2$ implies that two distinct geodesics in $\mathbb{H}^2$ above γ cannot have a common endpoint at infinity. This fact is needed in the proof of this and many later results.

Proof. As l and m both lie above γ, there is g in $\pi_1(M)$ such that $l = gm$. Let g denote any element of $\pi_1(M)$ such that $l = gm$, and let λ_g denote the axis of g in $\mathbb{H}^2$. Clearly, λ_g meets l if and only if λ_g meets m. As l and m are oppositely oriented, they must be disjoint from λ_g. Now consideration of the various possibilities shows that λ_g must lie in the region of $\mathbb{H}^2$ between l and m, and must not separate l from m. See Figure 3. In particular, it is now clear that m cannot meet $gl = g^2m$. This holds for any g such that $l = gm$. Thus m is also disjoint from $\alpha^k gl$, for each integer k, where α generates the

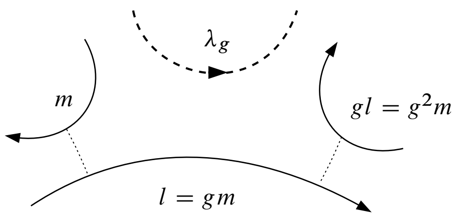


Figure 3. If the geodesics l and m are disjoint and oppositely oriented.

stabilizer of l. It follows immediately that the geodesics in the family $\{\alpha^k m, \alpha^k gl\}_{k\in\mathbb{Z}}$ are all disjoint, and disjoint from l. Now we consider the orthogonal projections onto l of these geodesics. These must all be disjoint, as any two of them can be separated by a geodesic perpendicular to l, constructed as in the proof of the previous lemma. The projection of gl to l is the translate by g of the projection of l to $g^{-1}l = m$, and so has the same length. We conclude that the orthogonal projections onto l of the geodesics in the family will all have the same length. Now it follows that this length can be at most $l(\gamma)/2$. Equality can only occur if there are distinct geodesics in the family $\{\alpha^k m, \alpha^k gl\}_{k\in\mathbb{Z}}$ with a common endpoint at infinity. This is not possible by Remark 2.4, so it follows that the orthogonal projection of m onto l has length strictly less than $l(\gamma)/2$, as required. ∎

Next we will obtain analogous results in the case of crossing geodesics, even though we can no longer compare orientations of such geodesics in the same way.

Lemma 2.5. *Let γ be a closed geodesic on an orientable hyperbolic surface M, and let l and m be crossing geodesics in $\mathbb{H}^2$ above γ. If the orthogonal projection of m to l has orientation opposite from that of l, then this projection has length strictly less than $l(\gamma)/2$.*

Proof. The orientation of γ gives orientations for l and m and their translates. Let l_- and l_+ be the points at infinity of l, and let m_- and m_+ be the points at infinity of m in the given directions.

Let α be the generator of the stabilizer of l, and let g be a covering translation that maps m to l, so g maps the point $p = m \cap l$ to the point $p' = l \cap gl$. We can assume (by composing g with some power of α if necessary) that p' lies between p and αp. See Figure 4 (a), where the positive ends of the lines correspond to the tips of the arrows.

We will show that $m = g^{-1}l$ and $m' = gl$ are disjoint. By hypothesis, the angle $m_- p l_+$, which is equal to the angle $l_- p' m'_+$, is less than $\pi/2$. Thus, the rays pm_+ and $p'm'_-$ cannot cross.

Now we will show that the rays pm_- and $p'm'_+$ also cannot cross. Let λ be the bisector of the angle $m_- p l_+$, so $g\lambda$ is the bisector of the angle $l_- p' m'_+$. Then λ and $g\lambda$ must be disjoint. Indeed, otherwise they would meet at a point x that is equidistant from p and p', and as g is orientation preserving and maps λ to $g\lambda$, sending p to p', it would have to fix x, contradicting the fact that g cannot have fixed points. As λ and $g\lambda$ are disjoint, they separate the rays pm_- and $p'm'_+$, so these rays cannot meet. Hence $m = g^{-1}l$ and $m' = gl$ are disjoint, as required.

Now the covering translation αg^{-1} maps $p' = m' \cap l$ to $\alpha p = l \cap \alpha m$, so it maps $m' = gl$ to αl, which is equal to l, and the preceding argument shows that m' is disjoint from αm. Thus, m and its translates by powers of α are disjoint from m' and its translates by powers of α. As the projection of m' to l has the same length as the projection of m to l (because m and m' make the same angle with l), and the sum of the two lengths is less than the translation length of α, it follows that the projection of m to l is shorter than $l(\gamma)/2$. Again we need Remark 2.4 to ensure this inequality is strict. ∎

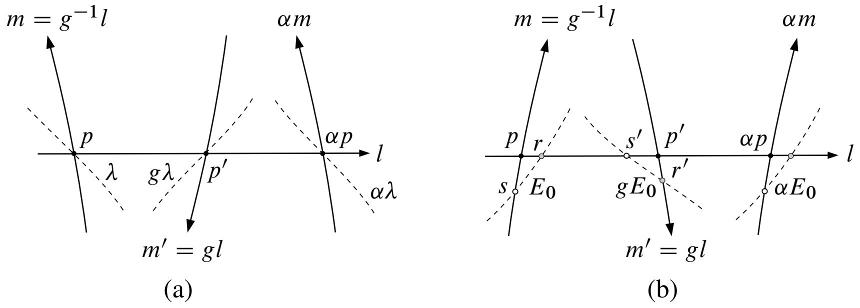


Figure 4. (a) Projection is oppositely oriented. (b) Overlap is oppositely oriented.

Lemma 2.6. *Let γ be a closed geodesic on an orientable hyperbolic surface M, and let l and m be crossing geodesics in $\mathbb{H}^2$ above γ whose images A and B in T overlap with opposite orientations. Then the orthogonal projection of m onto l has length strictly less than $l(\gamma)/2$.*

Proof. We will use the same notation as in the proof of the previous lemma. Let α be the generator of the stabilizer of l, and let g be a covering translation that maps m to l and sends the point $p = m \cap l$ to the point $p' = l \cap gl$ that lies between p and αp. We want to show that $m = g^{-1}l$ and $m' = gl$ are disjoint. Indeed, then the same argument as in the proof of Lemma 2.5 shows that the projection of m to l is shorter than $l(\gamma)/2$.

We can assume that the angle m_-pl_+, which is equal to the angle $l_-p'm'_+$, is greater than $\pi/2$, as otherwise the previous lemma gives the result. So the rays pm_- and $p'm'_+$ cannot cross. See Figure 4 (b).

It remains to show that the rays pm_+ and $p'm'_-$ cannot cross. By hypothesis, the axes A and B in T overlap with opposite orientations, so there is an edge e_0 of T that is traversed by A and B in different directions. If E_0 is the cutting geodesic dual to e_0, then l and m cross E_0 in opposite directions. Thus, gE_0 is also a cutting geodesic, and E_0 and gE_0 are disjoint (they cannot be equal because they cross l in opposite directions).

If E_0 crosses l and m at p, then gE_0 crosses l and m' at p' and the arcs E_0 and gE_0 play the roles of the bisectors λ and $g\lambda$ in Figure 4 (a). As E_0 and gE_0 are disjoint, they must separate the rays pm_+ and $p'm'_-$, which are therefore disjoint.

If E_0 crosses l and m at two different points r and s, then either r is in the ray pl_- and s is in the ray pm_+, or r is in the ray pl_+ and s is in the ray pm_-. See Figure 4 (b). In either case, as E_0 and gE_0 are disjoint, E_0 cannot cross m' and gE_0 cannot cross m, so E_0 and gE_0 must again separate the ends of the rays pm_+ and $p'm'_-$, hence these two rays must be disjoint.

The above argument shows that, in all cases, $m = g^{-1}l$ and $m' = gl$ are disjoint, as required. ∎

Lemma 2.7. *Let γ be a closed geodesic on an orientable hyperbolic surface M, and let l and m be crossing geodesics in $\mathbb{H}^2$ above γ. Then the orthogonal projections of the bisectors of l and m onto m have lengths whose sum is at most $2l(\gamma)$.*

Proof. Let α be a generator of the stabilizer of l, and let β be the corresponding generator of the stabilizer of m, so that they are conjugate in $\pi_1(M)$. Let λ and λ' denote the bisectors of l and m as shown in Figure 5. We will show that the geodesics λ, $\alpha\lambda'$ and $\alpha^2\lambda$ are disjoint, so their orthogonal projections to l are also all disjoint. This will imply that the sum of the orthogonal projections of λ and $\alpha\lambda'$ to l is at most the translation length of α^2, which is $2l(\gamma)$, as required.

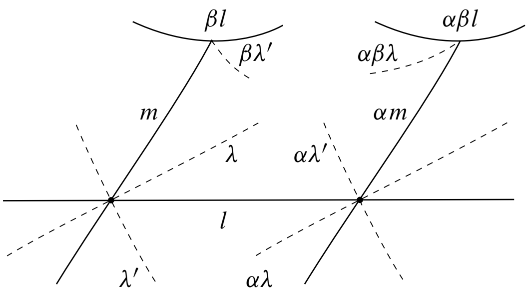


Figure 5. The bisectors of two crossing geodesics.

Let p be the image of $l \cap m$ in γ. Consider the commutator $\kappa = \alpha\beta\alpha^{-1}\beta^{-1}$, which must be a hyperbolic or parabolic isometry of $\mathbb{H}^2$. Thus κ is represented by a loop δ on M that follows γ turning left or right (but always to the same side) each time that it reaches p, so δ covers the image of γ four times before closing up. The loop δ is the projection of a piecewise geodesic line k in $\mathbb{H}^2$ formed by an arc of $\alpha^{-1}l$, an arc of m, an arc of l, an arc of αm and their translates by the action of $\kappa = \alpha\beta\alpha^{-1}\beta^{-1}$. See Figure 5. The bisectors at the corners of k are $\beta\lambda'$, λ, $\alpha\lambda'$, $\alpha\beta\lambda$ and their translates by powers of $\kappa = \alpha\beta\alpha^{-1}\beta^{-1}$.

By the symmetry of the picture, if λ were to intersect $\alpha\lambda'$, then $\beta\lambda'$ would intersect $\alpha\lambda'$ at the same point, and so all the bisectors at the corners of k would cross at that point, which would be fixed by $\kappa = \alpha\beta\alpha^{-1}\beta^{-1}$ and so κ would be elliptic, a contradiction. ∎

Now we start the proof of the main part of Theorem 5.1, which is the case when l and m cross and their images in T overlap with coherent orientations.

Let γ be a closed geodesic on a hyperbolic three-punctured sphere or once-punctured torus F, and let l and m be crossing geodesics in $\mathbb{H}^2$ above γ. As in the proof of Lemma 2.1, there is g in $\pi_1(M)$ such that $m = gl$, and l and m project in T to axes of α and $g\alpha g^{-1}$, respectively, which we denote by A and B.

Part (2) of Lemma 1.2 tells us that the intersection $A \cap B$ is a point or an edge path of T of length at most $L(\alpha) - 2$, where $L(\alpha)$ denotes the translation length of α acting on T. (Note that we cannot have $L(\alpha) = 1$, as this would imply that the closed geodesic γ

is simple, contradicting the fact that l and m cross.) Thus the intersection $A \cap \alpha B$ is a point or an edge path of the same length as $A \cap B$ and so is disjoint from $A \cap B$. In particular, there are two consecutive edges e and e' of A whose union meets each of $A \cap B$ and $A \cap \alpha B$ in at most one vertex and separates them in A. The cutting geodesics E and E' in $\mathbb{H}^2$ which correspond to e and e' must cross l, not cross m nor αm, and must separate m from αm. In particular, m and αm must be disjoint. But as m and αm cross l, it does not immediately follow that there is a geodesic in $\mathbb{H}^2$ which crosses l orthogonally and separates m from αm. The geodesics E and E' are edges of an ideal quadrilateral region Q in $\mathbb{H}^2$ such that Q is disjoint from m and αm. We call Q a gap quadrilateral. Note that there may be several gap quadrilaterals between m and αm. The preceding argument shows only that there must be at least one. Of course, l does meet Q in some arc. There are two possible configurations. One is that l crosses opposite edges of Q, and the other is that l crosses adjacent edges of Q. In the second case, the adjacent edges of Q have a common vertex v, i.e., a point in the circle at infinity of $\mathbb{H}^2$. We will say that l crosses Q at a cusp and that v is the associated vertex.

Recall that $\pi_1(F)$ is a free group of rank 2 and that $\pi_1(F)$ acts freely on the tree T dual to the cutting geodesics in $\mathbb{H}^2$ given by the edges of the quadrilaterals. These edges determine a new set of generators of $\pi_1(F)$, which we will denote by x and y, and α is uniquely expressible as a reduced word in x and y whose length we denote by $l(\alpha)$. Consider the usual projection of l and m to the axes A and B in the tree T. We let $L(\alpha)$ denote the length of the cyclically reduced word W conjugate to α, so that α acts on A by a translation of length $L(\alpha)$. Then the infinite reduced word $\widetilde{W}$ associated to A, the axis of α in T, is made by concatenating copies of W, and each letter of W corresponds to an oriented edge e of A. Further, each oriented edge of A corresponds to l crossing one of the cutting geodesics which are edges of our tiling of $\mathbb{H}^2$ by ideal quadrilaterals. Let w denote the subword of $\widetilde{W}$ associated to the interval $A \cap B$. We will call w the *overlap word of* l. Observe that the overlap word is empty or trivial if A and B do not intersect or if they intersect at a single point. If w is nontrivial, then we will usually cycle W so that w is an initial segment of W. The final segment of W will be called the *gap word of* l. Similarly, the conjugate α^g acts on B by a translation of length $L(\alpha)$, and the infinite reduced word associated to B is also equal to $\widetilde{W}$. Let W' denote the subword of length $L(\alpha)$ obtained by reading along B, whose initial segment is the word w associated to the interval $A \cap B$. The final segment of W' will be called the *gap word of* m. Note that the two gap words are usually very different. In particular, neither can start or end with the same letter as the other, as that would contradict the fact that w is associated to the full intersection $A \cap B$.

It will be convenient to introduce the following terminology. We will say that two reduced words in x, y and their inverses are *equivalent* if they are equal or become equal after possibly interchanging x and y and/or inverting x or y. Note that each of these operations is an automorphism of the free group $\pi_1(F)$ on x and y. The automorphisms of $\pi_1(F)$ generated by these automorphisms will be called *elementary*. Thus, if two words w and w' in x and y are equivalent, there is an elementary automorphism of

$\pi_1(F)$ which sends w to w'. Such an automorphism simply corresponds to a re-labelling of the generators of F.

At this point, our proof of Theorem 5.1 breaks up into several cases, depending on whether F is a three-punctured sphere or a once-punctured torus and on the configuration of l in the various gap quadrilaterals. We will use a combination of arguments in hyperbolic geometry and arguments with reduced words in a free group.

As the cases of the three-punctured sphere and once-punctured torus are substantially different, we will devote a separate section to each.

3. The case of the three-punctured sphere

Lemma 3.1. *Let γ be a closed geodesic on a hyperbolic three-punctured sphere F, and let l and m be geodesics in $\mathbb{H}^2$ above γ. If there is a gap quadrilateral Q such that l crosses opposite edges of Q, then the orthogonal projection of l onto m has length strictly less than $l(\gamma)$.*

Proof. Let E_1 and E_2 be the edges of Q crossed by l, and let λ and λ' be the diagonals of Q. So l crosses λ and λ', forming a triangle. As F is a three-punctured sphere, λ and λ' cross at right angles, and therefore the perpendicular μ to l through the point $\lambda \cap \lambda'$ goes through the shaded regions in Figure 6 (a), and so it lies in the region of $\mathbb{H}^2$ bounded by E_1 and E_2, separating E_1 from E_2. (This type of argument will be used several times later.) If α is a generator of the stabilizer of l, then m lies between μ and $\alpha\mu$ or between μ and $\alpha^{-1}\mu$, so the orthogonal projection of m onto l is shorter than the translation length of α, and so is shorter than $l(\gamma)$, as required. ∎

Lemma 3.2. *Let γ be a closed geodesic on a hyperbolic three-punctured sphere, and let l and m be crossing geodesics in $\mathbb{H}^2$ above γ whose images in T do not share an edge. Then the orthogonal projection of m onto l has length strictly less than $l(\gamma)$.*

Proof. As the images of l and m in T do not overlap, l and m cannot cross together any edges of the quadrilaterals. So, l and m cross at an interior point p of a quadrilateral Q, and l and m cross opposite edges of Q. Let α be a generator of the stabilizer of l and let g be an element of $\pi_1(F)$ that maps m to l, chosen so that gp lies between p and αp. Then gQ is a gap quadrilateral for l, and l crosses opposite edges of gQ, so by the previous lemma, the orthogonal projection of m to l is shorter than $l(\gamma)$. ∎

Next we apply this result.

Lemma 3.3. *Let γ be a closed geodesic on a hyperbolic three-punctured sphere, and let l and $m = gl$ be geodesics in $\mathbb{H}^2$ above γ. Then either the projection of m onto l has length strictly less than $l(\gamma)$, or l crosses all the gap quadrilaterals at cusps with the same vertex.*

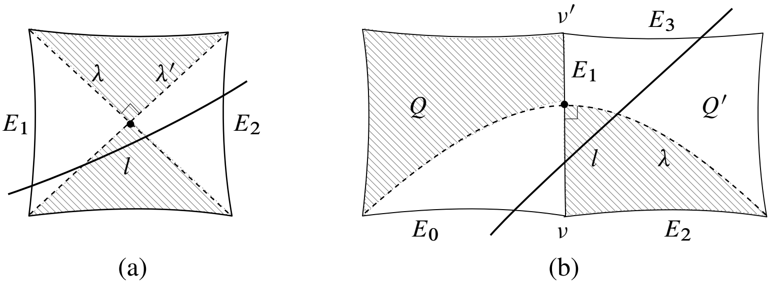


Figure 6. (a) The geodesic l crossing opposite sides of a gap quadrilateral. (b) The geodesic l crossing two gap quadrilaterals at cusps with different vertices.

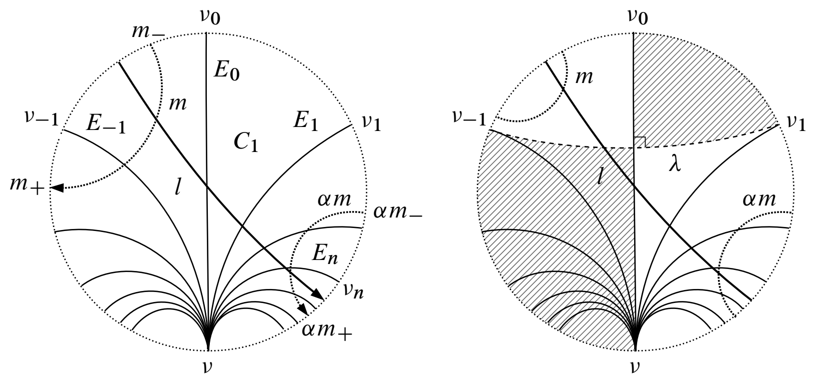
Proof. By Lemma 3.1, we can assume that l crosses each gap quadrilateral at a cusp. Either all these cusps have the same associated vertex, or there is a gap quadrilateral Q such that l crosses Q at a cusp with vertex v and crosses the next gap quadrilateral Q' at a cusp with a different vertex v'. See Figure 6 (b). Let E_1 be the edge that separates Q from Q', and let E_0 be the other edge of Q crossed by l. Let E_2 be the other edge of Q' with one endpoint at v, and let E_3 be the other edge of Q' crossed by l. Thus, m and all its translates by the action of the stabilizer of l lie outside the region of $\mathbb{H}^2$ bounded by E_0 and E_3.

As F is a three-punctured sphere, reflection in E_1 interchanges E_0 and E_2, so the hyperbolic line λ that joins the other endpoints of E_0 and E_2 is perpendicular to E_1. As l crosses from E_0 to E_3, it must cross the lines E_1 and λ, forming a right triangle (see Figure 6 (b)). Thus, the perpendicular to l from the point $E_1 \cap \lambda$ crosses l between $l \cap E_1$ and $l \cap \lambda$, and so it goes through the shaded regions in the picture. Hence this perpendicular is contained in the region of $\mathbb{H}^2$ bounded by E_0 and E_3 and separates E_0 from E_3. As m lies on the other side of E_0 and αm lies on the other side of E_3, this perpendicular separates m from αm. Thus, as before, the projection of m to l must be shorter than the translation length of α, which is $l(\gamma)$, as required. ■

We are left with the case where l crosses all the gap quadrilaterals at cusps with the same associated vertex.

Lemma 3.4. *Let γ be a closed geodesic on a hyperbolic three-punctured sphere F, and let l and $m = gl$ be crossing geodesics in $\mathbb{H}^2$ above γ, whose images in T overlap with coherent orientations. If l crosses all the gap quadrilaterals at cusps with the same vertex v, then m or αm do not cross any cusp edges at v.*

Proof. Label the cutting geodesics that end in v by E_i, $i \in \mathbb{Z}$, so that l crosses $E_0, E_1, \ldots, E_n$. We will call these E_i's cusp edges, and the regions between them cusp regions. Let Q be a gap quadrilateral, and let E_q and E_{q+1} be the edges of Q crossed by l. As Q separates m and αm, it follows that m can only cross E_i's with $i < q$, and αm can only



(a) (b)

Figure 7. The geodesic l crossing the cusps at a vertex.

cross E_i's with $i > q + 1$, as shown in Figure 7 (a). Note that if m crosses a cusp edge E_i, which is also crossed by l, then the axes A and B in T will share the edge of T corresponding to E_i. Thus the letter we read as l crosses E_i is part of the overlap word of l and m. A similar comment holds if αm crosses a cusp edge which is also crossed by l.

Now an orientation for γ induces orientations for l, m and αm. As m crosses l, the two endpoints m_+ and m_- of m lie on opposite sides of l. So αm has endpoints αm_+ and αm_-, and as α preserves orientation, m_+ and αm_+ lie on one side of l and m_- and αm_- lie on the other side. This implies that m and αm "travel around v in opposite directions", as shown in Figure 7 (a).

Now suppose that both m and αm cross at least one E_i. It follows that m and αm cross the E_i's in opposite directions. Hence one of m and αm crosses the E_i's in the opposite direction to l. Suppose that m does this, as shown in Figure 7 (a). (If we reverse the orientations of m and αm in the figure, then αm will do this.)

Recall that F has three cusps, and that the simple loops around these cusps represent x, y and xy, respectively, when correctly oriented. Thus, by an elementary automorphism of $\pi_1(F)$, we can arrange that as l crosses $E_0, E_1, \ldots, E_n$, either each crossing contributes x to the associated word, or each crossing contributes x and y alternately. In particular, the gap word for l is positive. As the words W and W' associated to l and m are conjugates and are reduced, the gap words for l and m have the same abelianization and are reduced. In particular, the gap word for m must also be positive. As m crosses the E_i's in the opposite direction to l, these crossings yield a negative word, which must therefore be disjoint from the gap word for m. But this implies that these crossings yield part of the overlap word, which is also impossible, as the overlaps of A and B in T are coherently oriented. This contradiction shows that m cannot cross any cusp edges at v, as required. ∎

Lemma 3.5. *Let γ be a closed geodesic on a hyperbolic three-punctured sphere, and let l and m be crossing geodesics in $\mathbb{H}^2$ above γ whose images in T overlap with coherent orientations. If l crosses all the gap quadrilaterals at cusps with the same vertex v, then the orthogonal projection of l onto m has length strictly less than $l(\gamma)$.*

Proof. Let $E_0, E_1, \ldots, E_n$ be the cusp edges crossed by l. By Lemma 3.4, either m or αm does not cross any E_i's, so by interchanging the roles of m and αm if necessary, we can assume that m does not cross any E_i's, so that the gap quadrilaterals start at E_0. As there is at least one gap quadrilateral, αm does not cross any E_i with $i \leq 1$. Let v_i be the endpoint of E_i other than v. Recall that the union of all the cusp edges that end at v is symmetric under reflection in E_0. In that reflection, the point v_1 is sent to v_{-1}. Let λ denote the geodesic joining v_1 to v_{-1}, see Figure 7 (b). The symmetry implies that λ meets E_0 orthogonally. As l crosses E_0 and E_1, it must also cross λ. As E_0 and λ form a right triangle with l, the perpendicular to l from the point $E_0 \cap \lambda$ crosses l between $l \cap E_0$ and $l \cap \lambda$. So this perpendicular to l has one endpoint between v_0 and v_1 and the other endpoint between v and v_{-1}. As m does not meet any E_i, this perpendicular does not cross m nor αm and it separates m from αm. Now the usual argument implies that the projection of m to l is shorter than $l(\gamma)$. ∎

Lemmas 2.1, 2.6, 3.2, 3.3 and 3.5 together show that Theorem 5.1 holds when M is the three-punctured sphere.

4. The case of the once-punctured torus

In this section, we will proceed much as in the previous one, but the proof is more delicate. The basic reason for this is that, unlike the three-punctured sphere, the once-punctured torus is not rigid, but admits a 2-parameter family of complete hyperbolic metrics, and when we cut a once-punctured torus into a hyperbolic quadrilateral, the resulting tessellation of $\mathbb{H}^2$ may be much less symmetric than the tessellation from the three-punctured sphere: the only symmetries may be the covering translations. In particular, the diagonals of the quadrilaterals do not have to intersect at right angles, and the family of cusp edges $\ldots, E_{-2}, E_{-1}, E_0, E_1, E_2, \ldots$ that end at a vertex v may not be invariant under reflections in those edges. We need to discuss what symmetries still exist.

We claim that if one fixes v and restricts attention to just the even-numbered E_i's, this family is invariant under reflection in any one of them, and the same holds for the family of odd-numbered edges. To see this, refer to Figure 8. The even-numbered cusp edges project to a single cutting geodesic λ_x in the punctured torus F, and the odd-numbered cusp edges project to a single cutting geodesic λ_y in F. There is a simple closed geodesic c_x representing x which meets λ_x in one point, and there is a simple closed geodesic c_y representing y which meets λ_y in one point. The punctured torus F admits an orientation preserving symmetry ρ of order two with three fixed points, the Weierstrass rotation. One

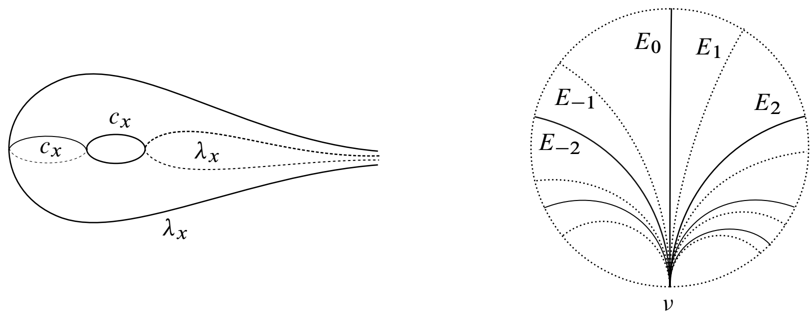


Figure 8. Symmetries of the cusp edges for a punctured torus.

fixed point is $c_x \cap c_y$, one is $c_x \cap \lambda_x$, and one is $c_y \cap \lambda_y$. Thus each of c_x, c_y, λ_x and λ_y is preserved but reversed by ρ. It follows that the family of even-numbered E_i's is invariant under reflection in any of them, as claimed. Similarly, the same holds for the family of odd-numbered edges. Note that if λ and λ' have a common endpoint v, then the bisector of λ and λ' must also share that endpoint. It follows from the above that the reflections in the bisectors of two consecutive E_i's preserve the family of all E_i's, interchanging the even-numbered E_i's with the odd-numbered E_i's. Note that these reflections in the E_i's and in the bisectors need not come from a symmetry of the once-punctured torus, and they may not preserve the quadrilaterals of the tessellation.

We will continue to use the terminology introduced at the end of Section 2. By Lemmas 2.1 and 2.6, we are left only with the case when l and m are crossing geodesics whose axes in T do not overlap or overlap with coherent orientations.

Lemma 4.1. *Let γ be a closed geodesic on a hyperbolic once-punctured torus F, and let l and $m = gl$ be geodesics in $\mathbb{H}^2$ above γ whose images in T overlap with coherent orientations. Then l crosses at least one gap quadrilateral at a cusp.*

Proof. If l does not cross any gap quadrilateral at a cusp, then each edge of each gap quadrilateral crossed by l yields the same letter in $\widetilde{W}$, say x, so that the gap word of l is a positive power of x.

Let w denote the overlap subword of $\widetilde{W}$ associated to the interval $A \cap B$. Then, by cycling W if needed, we can assume that $W = wx^n$. Reading along the axis B yields that, after cycling, $W' = wz$ for some word z that is the gap word for m. As W' is conjugate to W, they have the same image in the abelianization of the free group $\pi_1(F)$. Thus $z = x^n$, but then there is no gap. This contradiction completes the proof of the lemma. ∎

Next, we consider some other special cases.

Lemma 4.2. *Let γ be a closed geodesic on a hyperbolic once-punctured torus F, and let l and $m = gl$ be geodesics in $\mathbb{H}^2$ above γ whose images in T overlap with coherent orientations. If l has only one gap quadrilateral, then l and m are disjoint.*

Proof. Lemma 4.1 implies that l crosses the gap quadrilateral at a cusp. By applying an elementary automorphism of $\pi_1(F)$, if needed, we can assume that the two edges of the gap quadrilateral crossed by l yield the gap word xy in $\widetilde{W}$. Let w denote the overlap subword of $\widetilde{W}$ associated to the interval $A \cap B$. Then, by cycling W if needed, we can assume that $W = wxy$. Reading along the axis B yields $W' = wz$ for some word z. As W' is conjugate to W, they have the same image in the abelianization of the free group $\pi_1(F)$. Thus the word z must be xy or yx. In the first case, W' would equal W, contradicting the fact that w is the entire overlap word. It follows that $W' = wyx$. Now Figure 9 (a) shows that the two ends of m must lie on the same side of l, so l and m must be disjoint. ■

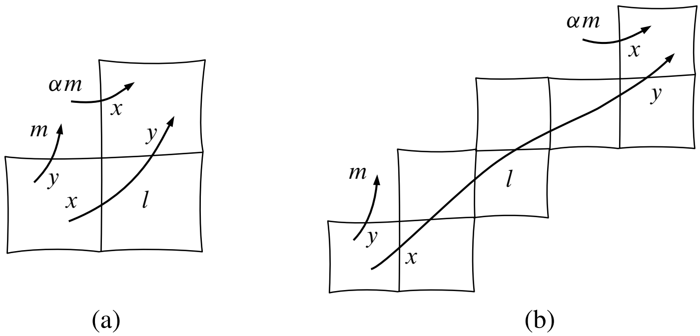


Figure 9. If the gap word is positive, starting in x and ending in y.

The argument above can be greatly generalized to obtain the following result.

Lemma 4.3. *Let γ be a closed geodesic on a hyperbolic once-punctured torus F, and let l and $m = gl$ be geodesics in $\mathbb{H}^2$ above γ whose images in T overlap with coherent orientations. If the gap word of l is a positive word in x and y that starts with x and ends with y, then l and m are disjoint, so the orthogonal projection of m onto l has length strictly less than $l(\gamma)$.*

Proof. As in the preceding lemma, if w denotes the overlap subword of $\widetilde{W}$ associated to the interval $A \cap B$, then $W = wxuy$ where u is a positive word, and $W' = wz$ for some word z. As W' is conjugate to W, they have the same image in the abelianization of the free group $\pi_1(F)$. Thus the word z is also positive. Further z cannot begin with x, nor end with y, as either would contradict the fact that w is the entire overlap word. It follows that $W' = wyu'x$, for some positive word u'. Now Figure 9 (b) shows that the two ends of m must lie on the same side of l, so that l and m must be disjoint. Thus Lemma 2.1 implies the required result. ■

Next we consider some special cases for subwords of gap words.

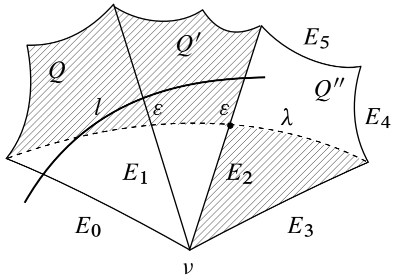


Figure 10. If the gap word contains $yxy^{-1}x$ or $yxy^{-1}y^{-1}$.

Lemma 4.4. *Let γ be a closed geodesic on a hyperbolic once-punctured torus, and let l and $m = gl$ be two geodesics in $\mathbb{H}^2$ above γ. If the gap word of l contains a subword $yxy^{-1}x$ or $yxy^{-1}y^{-1}$, then the projection of m to l is shorter than γ.*

Proof. By hypothesis, l crosses two gap quadrilaterals Q and Q' at cusps with the same vertex v and crosses the next gap quadrilateral Q'' either at a cusp with a different vertex or across opposite edges. See Figure 10. Let E_0, E_1 and E_2 be the edges of Q and Q' crossed by l and with endpoint v, let E_3 be the other edge of Q'' with one endpoint at v, and let E_4 and E_5 be the other edges of Q''. So m and all its translates by the action of the stabilizer of l lie outside the region of $\mathbb{H}^2$ bounded by E_0, E_4 and E_5. Let λ be the hyperbolic line that joins the other endpoints of E_0 and E_3.

As the reflection in the bisector of E_1 and E_2 interchanges E_0 and E_3, the angles that λ makes with E_1 and E_2 inside Q' (marked in the picture by ε) are equal, and so ε is larger than $\pi/2$. As l crosses from E_0 to E_4 or E_5, it crosses the lines E_2 and λ, forming a triangle whose other angles must be acute. Therefore, the perpendicular to l from the point $E_2 \cap \lambda$ crosses l between $l \cap \lambda$ and $l \cap E_2$, so it goes through the shaded regions in the picture. Hence it lies inside the region of $\mathbb{H}^2$ bounded by E_0, E_4 and E_5, and separates E_0 from E_4 and E_5. As m lies on the other side of E_0, and αm lies on the other side of E_4 or E_5, this perpendicular must separate m from αm, so as before, the projection of m to l must be shorter than the translation length of α, which is $l(\gamma)$. ∎

Lemma 4.5. *Let γ be a closed geodesic on a hyperbolic once-punctured torus, and let l and $m = gl$ be two geodesics in $\mathbb{H}^2$ above γ. If the gap word of l contains a subword $yx^k y^{-1}$, with $k > 1$, then the projection of m to l is shorter than γ.*

Proof. By hypothesis, l crosses a gap quadrilateral Q_0 at a cusp with vertex v, then crosses $k-1$ gap quadrilaterals $Q_1, Q_2, \ldots, Q_{k-1}$ across opposite edges and then crosses another gap quadrilateral Q_k at a cusp with vertex $v' \neq v$, as in Figure 11.

Let E_0 and E_1 be the edges of Q_0 crossed by l, with E_1 separating Q_0 from Q_1, and let E_{k-1} and E_k be the edges of Q_k crossed by l, with E_{k-1} separating Q_{k-1} from Q_k. Let τ be the covering translation that takes Q_0 to Q_1. Note that τ must also take Q_{i-1}

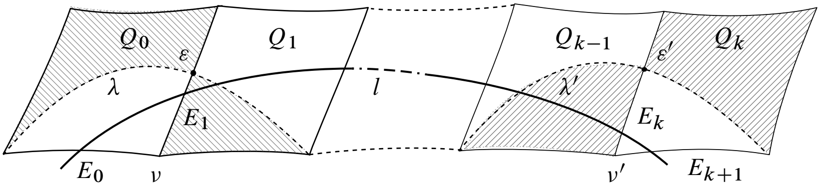


Figure 11. If the gap word contains yx^ky^{-1}, $k > 1$.

to Q_i, for $2 \le i \le k$. In particular, τ^{k-1} takes E_1 to E_k. Let λ be the geodesic joining $\tau^{-1}v$ to τv, and let λ' be the geodesic joining $\tau^{-1}v'$ to $\tau v'$. Thus τ^{k-1} also takes λ to λ'. Also l must cross λ and λ'. Therefore, the angles formed by E_1 and λ, and by E_k and λ' are equal, and so the angles marked in Figure 11 by ε and ε' are supplementary. If $\varepsilon \ge \pi/2$, the perpendicular from the point $\lambda \cap E_1$ to l goes through the shaded regions between λ and E_1, and if $\varepsilon' \ge \pi/2$, the perpendicular from the point $\lambda' \cap E_k$ to l goes through the shaded regions between λ' and E_k. So one of these two perpendiculars lies in the region of $\mathbb{H}^2$ bounded by E_0 and E_{k+1} and separates E_0 from E_{k+1}, thus separating m from αm. As before, this implies that the projection from m to l must be shorter than $l(\gamma)$. ■

Lemma 4.6. *Let γ be a closed geodesic on a hyperbolic once-punctured torus, and let l and $m = gl$ be two geodesics in $\mathbb{H}^2$ above γ. If the gap word of l contains the subwords x^2 and y^2, then the projection of m to l has length less than γ.*

Proof. If the gap word for l contains the subwords x^2 and y^2, then l crosses a gap quadrilateral Q "from left to right" and crosses another gap quadrilateral Q' "from the bottom to the top", crossing both diagonals of each quadrilateral, see Figure 12. There is a covering translation τ that takes Q to Q' and takes the diagonals λ and λ' of Q to the diagonals of Q', so the angles between their diagonals are equal. Hence either $\varepsilon \ge \pi/2$ or $\varepsilon' \ge \pi/2$.

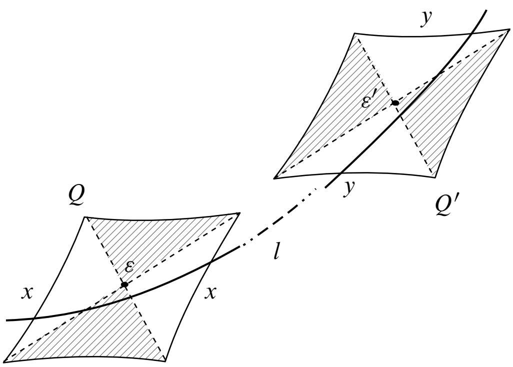


Figure 12. If the gap word contains the subwords x^2 and y^2.

In the first case, the perpendicular to l from $\lambda \cap \lambda'$ goes through the shaded regions of Q, and in the second case, the perpendicular to l from $\tau\lambda \cap \tau\lambda'$ goes through the shaded regions of Q'. So one of these perpendiculars to l lies in a region of $\mathbb{H}^2$ that is not crossed by m or αm, and so it separates m from αm. As before, this implies that the projection of m to l must be shorter than $l(\gamma)$. ■

Lemma 4.7. *Let γ be a closed geodesic on a hyperbolic once-punctured torus, and let l and m be crossing geodesics in $\mathbb{H}^2$ above γ whose images A and B in T do not overlap. Then the orthogonal projection of m onto l has length strictly less than $l(\gamma)$.*

Proof. As the axes A and B in T do not overlap, l and m cross in the interior of a quadrilateral Q, and l and m cross opposite edges of Q. Therefore, the word for l contains a subword $x^{\pm 2}$ and the word for m contains a subword $y^{\pm 2}$, or vice versa. As the words for l and m are conjugate, each must contain the other subword too. The quadrilaterals Q' and Q'' crossed by l and m corresponding to these other subwords are gap quadrilaterals for l and m, respectively.

Now the proof of the previous lemma shows that either there is a perpendicular to l in the region of $\mathbb{H}^2$ between the edges of Q' crossed by l, or there is a perpendicular to m in the region of $\mathbb{H}^2$ between the edges of Q'' crossed by m. As before, this implies either that the orthogonal projection of m to l is shorter than $l(\gamma)$ or that the orthogonal projection of m to l is shorter than $l(\gamma)$. ■

By using the above lemmas, we will show how to reduce to the case when the gap word has a very special form.

Lemma 4.8. *Let γ be a closed geodesic on a hyperbolic once-punctured torus F, and let l and $m = gl$ be geodesics in $\mathbb{H}^2$ above γ whose images in T overlap with coherent orientations. By applying elementary automorphisms of $\pi_1(F)$ and replacing α by its inverse if needed, we can arrange that one of the following cases holds:*

(1) *The projection of m to l has length less than γ.*

(2) *The gap word of l is positive and of the form $yx^{d_1}yx^{d_2}y \dots yx^{d_r}y$ or the form $x^{d_1}yx^{d_2}y \dots yx^{d_r}$, where each $d_i \geq 1$.*

(3) *The gap word of l is a subword of $(xyx^{-1}y^{-1})^N$ for some N.*

Remark 4.9. In part (3), the geodesic l crosses all the gap quadrilaterals at cusps with the same associated vertex.

Proof. Suppose that the projection of m to l has length greater than that of γ. We will show that part (2) or (3) of the lemma must hold.

Lemma 4.6 tells us that the gap word v of l cannot contain both x^2 and y^2. By applying elementary automorphisms, it also follows that v cannot contain both x^2 and y^{-2}, nor both x^{-2} and y^2, nor both x^{-2} and y^{-2}. It follows that one of x and y can only occur with exponents 1 or -1. We will assume that y can only occur with exponents 1 or -1.

Suppose that x has a power which is not 1 or -1. After an elementary automorphism, we can assume that v contains x^k for some $k \geq 2$. The proof of Lemma 4.1 shows that v cannot equal a power of x, so by a further elementary automorphism, and perhaps the inversion of α, we can assume that v contains the subword $x^k y$.

If v contains the subword $x^k y x^l$, we claim that $l \geq 1$. Indeed, if $l \leq -1$, the gap word v contains the subword $x^2 y x^{-1}$, which is equivalent to the inverse of yxy^{-2}. Now Lemma 4.4 shows that this cannot occur.

If v contains the subword $x^k y x^l y^m$, we claim that $m = 1$. For if $l \geq 2$ and $m = -1$, then v contains the subword $yx^l y^{-1}$, which is not possible by Lemma 4.5. And if $l = 1$ and $m = -1$, then v contains the subword $xyxy^{-1}$ which is equivalent to the inverse of $yxy^{-1}x$, and so impossible by Lemma 4.4.

If v contains the subword $x^k y x^l y x^n$, we claim that $n \geq 1$. If $l \geq 2$, the same argument we used above shows that $n \geq 1$. If $l = 1$ and $n \leq -1$, then v contains the subword $yxyx^{-1}$, which is equivalent to the inverse of $yxy^{-1}x$, and so impossible by Lemma 4.4.

Now a simple inductive argument shows that the entire segment of v which begins at x^k must be positive. By applying the same argument to the inverse of α, we conclude that we can arrange that the entire gap word v is positive.

Next, suppose that x also can only occur with exponents 1 or -1. Lemma 4.4 tells us that v cannot contain the subword $yxy^{-1}x$, nor any word which is equivalent to $yxy^{-1}x$ or its inverse. It follows that in the gap word v, either all powers of x have the same sign, and the same holds for all powers of y, or that v is a subword of $(xyx^{-1}y^{-1})^N$ for some N. In the first case, we can apply an elementary automorphism to arrange that v is positive. The second case is part (3) of the statement of the lemma.

Finally, we recall that if the gap word v of l is positive, so is that of m. It follows from Lemma 4.3 that v must begin and end with x or begin and end with y, which completes the proof of the lemma. ∎

At this point, we need to notice that the form of the gap words depends on our initial choice of generators x and y for $\pi_1(F)$. We will now discuss how to change generators so as to simplify the gap words we are considering.

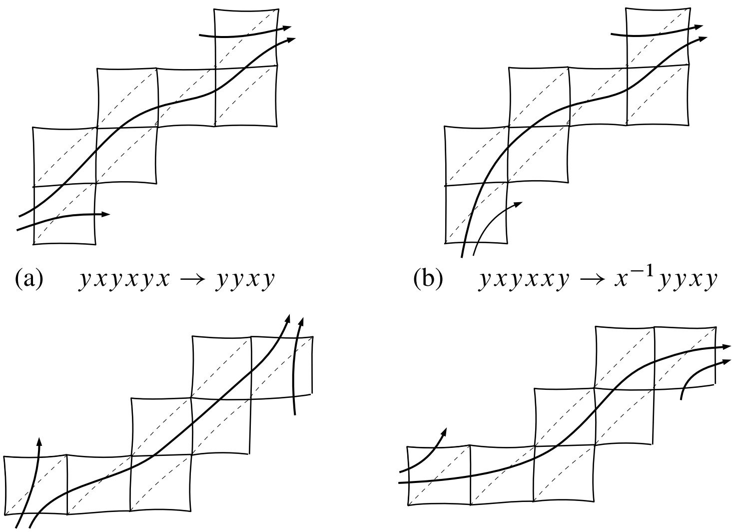
The basic operation is to replace one pair of opposite edges of a quadrilateral by diagonals. We will say that a reduced word in x and y is of *mixed sign* if it contains positive and negative powers of x, or if it contains positive and negative powers of y. Clearly, a reduced word of mixed sign cannot be equivalent to a positive word.

Lemma 4.10. *Let γ be a closed geodesic on a hyperbolic once-punctured torus F, and let l and $m = gl$ be geodesics in $\mathbb{H}^2$ above γ whose images in T overlap with coherent orientations. If the gap word of l is $yx^{d_1}yx^{d_2}y \ldots yx^{d_r}y$ or $x^{d_1}yx^{d_2}y \ldots yx^{d_r}$, where each $d_i \geq 1$, then the above change of basis yields a new gap word which either is of mixed sign or is positive and shorter than the original gap word.*

Proof. We start by noting that the gap words for l and m must begin and end in distinct letters. As we are assuming that the gap word for l is positive, it follows that the gap word for m is also positive. In particular, each gap word must begin and end in x or in y. As W and W' are cyclically reduced, it follows that the overlap word w cannot begin or end in x^{-1} or in y^{-1}.

Now replace the sides of the quadrilaterals corresponding to x by the diagonals shown dotted in Figure 13, keeping the names of the generators.

If the gap word is $yx^{d_1}yx^{d_2}y \dots yx^{d_r}y$ and the overlap word w starts and ends with x, then the new gap word is obtained from the original by reducing each d_i by 1. See Figure 13 (a). Thus the new gap word is positive and shorter, as required. If the overlap word starts or ends with y, the new gap word is of mixed sign. See Figure 13 (b).



(a) $yxyxyx \to yyxy$ (b) $yxyxxy \to x^{-1}yyxy$

(c) $xxyxyx \to xyy$ (d) $xxyxyx \to xxyyx$

Figure 13. A change of basis.

If the gap word is $x^{d_1}yx^{d_2}y \dots yx^{d_r}$, then the new gap word is obtained from the original by reducing each d_i by 1, for $1 < i < r$, while d_1 and d_r are either reduced by 1 or remain the same, depending on how the overlap word starts and ends. See Figures 13 (c), (d). Thus the new gap word is positive and is shorter, except possibly when $r = 2$. But this case cannot occur because it would mean that the gap word v of l equals $x^{d_1}yx^{d_2}$, so that the gap word v' of m contains only one y. Hence v' would have to begin or end in x, contradicting the fact that v and v' cannot begin or end with the same letter. ∎

Now we need to put together everything we have proved so far.

Lemma 4.11. *Let γ be a closed geodesic on a hyperbolic once-punctured torus F, and let l and $m = gl$ be geodesics in $\mathbb{H}^2$ above γ whose images in T overlap with coherent orientations. Then either the projection of m onto l has length strictly less than $l(\gamma)$, or l crosses all the gap quadrilaterals at cusps with the same associated vertex.*

Proof. Lemma 4.8 and Remark 4.9 together tell us that either the projection of m onto l has length strictly less than $l(\gamma)$, or the gap word of l is positive and has one of the forms $yx^{d_1}yx^{d_2}y\dots yx^{d_r}y$ or $x^{d_1}yx^{d_2}y\dots yx^{d_r}$, where each $d_i \geq 1$, or l crosses all the gap quadrilaterals at cusps with the same associated vertex, so it remains to handle the middle case.

Lemma 4.10 tells us that in this case, there is a change of basis yielding a new gap word which is either of mixed sign or positive and shorter than the original gap word. If the new gap word is positive and one of x and y only occurs with exponent 1, we can apply Lemma 4.10 again. Thus, by repeatedly applying this lemma, we must eventually obtain a gap word of mixed sign or a positive word that contains both x^2 and y^2. In the second case, Lemma 4.6 implies that the projection of m onto l has length strictly less than $l(\gamma)$. If the gap word v is of mixed sign, it cannot be equivalent to a positive word. Thus, Lemma 4.8 implies that either the projection of m onto l has length strictly less than $l(\gamma)$ or l crosses all the gap quadrilaterals at cusps with the same associated vertex, thus completing the proof of the lemma. ∎

Lemmas 2.6, 4.7 and 4.11 show that we can reduce to the case when the crossing geodesics l and m have images in T which overlap with coherent orientations, and l crosses all the gap quadrilaterals at cusps with the same vertex v. As in the proof of Lemma 3.4, we label the cusp edges at v by E_i, $i \in \mathbb{Z}$, so that $E_0, E_1, \dots, E_n$ are the cusp edges crossed by l, see Figure 14.

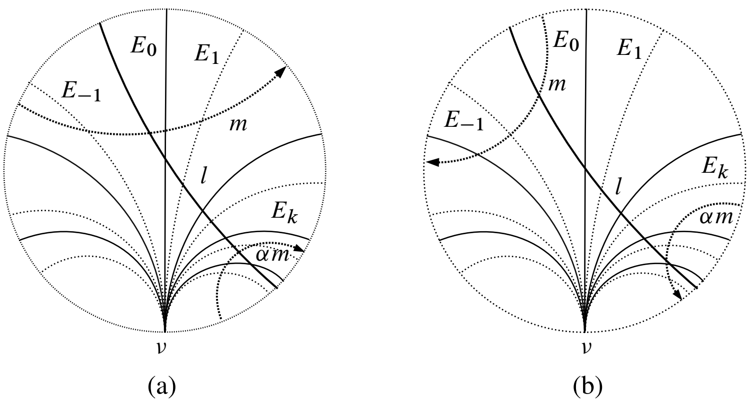


Figure 14. The cusps at a vertex for a punctured torus.

Lemma 4.12. *Let γ be a closed geodesic on a hyperbolic once-punctured torus F, and let l and $m = gl$ be crossing geodesics in $\mathbb{H}^2$ above γ whose images in T overlap with coherent orientations. If l crosses the k gap quadrilaterals at cusps with the same vertex v, then*

(1) *One of m or αm does not cross any of the cusp edges $E_0, E_1, \dots, E_n$.*

(2) *Each of m and αm crosses less than $k - 1$ of the remaining cusp edges.*

Proof. (1) As in the proof of Lemma 3.4, let Q be a gap quadrilateral, and let E_q and E_{q+1} be the edges of Q crossed by l. As Q separates m and αm, it follows that m can only cross E_i's with $i < q$, and αm can only cross E_i's with $i > q+1$, as shown in Figure 14 (a). From that same proof, we also know that m and αm "travel around v in opposite directions", as shown in Figure 14 (a).

Observe that if m crosses two E_i's, then it crosses all the E_j's between them.

Now suppose that m crosses E_0 and E_{-1}. This implies that the overlap between l and m starts as they cross E_0, so that l must cross all the overlap quadrilaterals and all the gap quadrilaterals at cusps with vertex v. But this implies that α fixes v and so is a parabolic element, contradicting our assumption that γ is a closed geodesic. We conclude that m cannot cross E_0 and E_{-1}. A similar argument shows that αm cannot cross E_n and E_{n+1}. Moreover, if m crosses E_0 then αm cannot cross E_n and vice versa, because m and αm cross the cusp edges in opposite directions, but l crosses E_0 and E_n in the same direction, and αm is a translate of m.

It follows that either m crosses only E_i's with $i < 0$, or αm crosses only E_i's with $i > n$, which proves the first part of the lemma.

(2) To prove this, we need only consider m, as the roles of m and αm can be interchanged. If m crosses some E_i with $i \geq 0$, then m cannot cross any E_i's with $i < 0$. Hence m can only cross E_i's with $0 \leq i < q$, which proves part (2) of the lemma in this case.

Now suppose that m crosses some E_i with $i < 0$, so that m cannot cross any E_i's with $i \geq 0$. Then m must cross E_{-1}, as otherwise E_{-1} would separate m from l. So the k gap quadrilaterals must start at E_0 and end at E_k, and if m crosses $k-1$ cusp edges, they must be $E_{-1}, E_{-2}, \ldots, E_{-k+1}$.

If w denotes the overlap word for l and m, then we can read off the rest of the word W from the cusp edges that l crosses in the gap, namely $E_0, \ldots, E_k$. As our punctured torus has only one cusp, after an elementary automorphism of $\pi_1(F)$, we can arrange that as l crosses these cusp edges, we read off the letters $x, y, x^{-1}, y^{-1}, x, y, x^{-1}, y^{-1}, \ldots$ As the images of l and m in T overlap and are oriented coherently, m crosses the cusp edges in the opposite direction to l. Thus as m crosses the cusp edges $E_{-1}, E_{-2}, \ldots, E_{-k+1}$, we read off the letters y, x, y^{-1}, x^{-1} repeatedly. This leads to four cases depending on the size of k modulo 4.

Case $k = 4N$, where $N \geq 1$: Then $W = w(xyx^{-1}y^{-1})^N x$. Hence we find that W' has initial segment $w(yxy^{-1}x^{-1})^{N-1}yxy^{-1}$. As $l(W') = l(W)$, we must have $W' = w(yxy^{-1}x^{-1})^{N-1}yxy^{-1}zu$, where each of z and u denotes one of x, y, x^{-1} or y^{-1}. As W' is a conjugate of W, abelianizing shows immediately that zu has weight 0 in x and in y. But this is impossible, as W', and hence zu, is a reduced word.

Case $k = 4N+1$, where $N \geq 0$: As in the preceding case, we have that $W = w(xyx^{-1}y^{-1})^N xy$, and W' has initial segment $w(yxy^{-1}x^{-1})^N$. It follows that $W' = w(yxy^{-1}x^{-1})^N zu$, where each of z and u denotes one of x or y or its inverses. As W' is a conjugate of W, abelianizing shows that zu must be equal to xy or to yx. The first case is impossible as W' is reduced, so we must have $zu = yx$. Thus we can write W in the form $wxUy$ and can write W' in the form $wyU'x$. Now the argument in the proof of

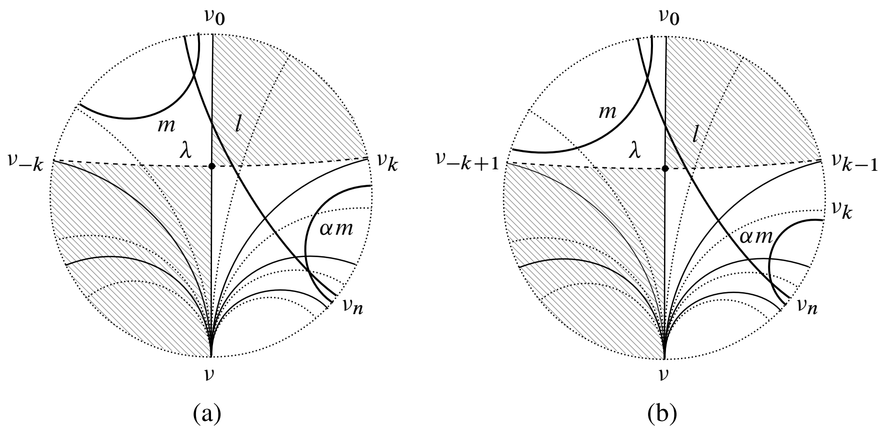


Figure 15. (a) If k is even. (b) If k is odd.

Lemma 4.3 shows that l and m must be disjoint, contradicting our assumption that they cross. Note that the use of Figure 9 (b) did not depend on the gap word xUy being positive. That hypothesis was used in the proof of Lemma 4.3 to show that the gap word for m must be of the form $yU'x$, but in the present situation, that is given.

Case $k = 4N + 2$, where $N \geq 0$: As in the case before, $W = w(xyx^{-1}y^{-1})^N xyx^{-1}$, and W' has initial segment $w(yxy^{-1}x^{-1})^N y$. Thus $W' = w(yxy^{-1}x^{-1})^N yzu$, where each of z and u denotes one of x or y or its inverses. As W' is a conjugate of W, abelianizing shows immediately that zu has weight 0 in x and in y. But this is impossible, as W', and hence zu, is a reduced word.

Case $k = 4N + 3$, where $N \geq 0$: We have that $W = w(xyx^{-1}y^{-1})^N xyx^{-1}y^{-1}$, and W' has initial segment $w(yxy^{-1}x^{-1})^N yx$ as in the preceding case. Thus $W' = w(yxy^{-1}x^{-1})^N yxzu$, where each of z and u denotes one of x or y or its inverses. As W' is a conjugate of W, abelianizing shows that zu must be equal to $x^{-1}y^{-1}$ or to $y^{-1}x^{-1}$. The first case is impossible as W' is reduced, so we must have $zu = y^{-1}x^{-1}$. Thus we can write W in the form $wxUy^{-1}$ and can write W' in the form $wyU'x^{-1}$. Now a similar argument to that in the proof of Lemma 4.3 shows that l and m must be disjoint, contradicting our assumption that they cross. Note that in this case, we need a modified version of Figure 9 (b). The above contradictions, for any value of k, complete the proof of part (2). ∎

Lemma 4.13. *Let γ be a closed geodesic on a hyperbolic once-punctured torus F, and let l and $m = gl$ be geodesics in $\mathbb{H}^2$ above γ whose images in T overlap with coherent orientations. If l crosses all the gap quadrilaterals at cusps with the same vertex, then the orthogonal projection of l onto m has length strictly less than $l(\gamma)$.*

Proof. Let $E_0, E_1, \ldots, E_n$ be the cusp edges crossed by l. By Lemma 4.12, we can assume that the k gap quadrilaterals start at E_0 and end at E_k, so αm can only cross E_i's with $i > k$, and m can only cross E_i's with $-(k-1) < i < 0$. Recall that the union of all the even cusp edges that end at v is symmetric under reflection in E_0. If v_i is the endpoint of the cusp edge E_i, this reflection sends the point v_{2i} to v_{-2i}, see Figure 15.

If k is even, we let λ denote the geodesic joining v_k to its reflected image v_{-k}. The symmetry implies that λ meets E_0 orthogonally. As l crosses the cusp geodesic E_k, it must also cross λ. Let μ denote the perpendicular to l from the point $E_0 \cap \lambda$. Clearly, one endpoint of μ lies between v_0 and v_k, and the other endpoint lies between v and v_{-k}. As m cannot cross E_{-k}, it follows immediately that μ cannot meet m nor αm. As usual, this implies that the orthogonal projection of l onto m has length strictly less than $l(\gamma)$.

If k is odd, we apply the above paragraph using the even number $k-1$ in place of k. As m cannot cross E_{-k+1}, it again follows immediately that μ cannot meet m nor αm, which again implies that the orthogonal projection of l onto m has length strictly less than $l(\gamma)$. Thus the result follows in either case, as required. ∎

Lemmas 2.1, 2.6, 4.7, 4.11 and 4.13 together show that Theorem 5.1 holds when M is the three-punctured torus.

5. The main result

At this point, we are ready to complete the proof of our main result.

Theorem 5.1. *Let γ be a closed geodesic on an orientable hyperbolic surface M, and let l and m be distinct geodesics in $\mathbb{H}^2$ above γ. Then the orthogonal projection of l onto m is shorter than the length of γ.*

Proof. The results in the preceding three sections yield a proof of the theorem in the special cases when M is finitely covered by a three-punctured sphere or by a once-punctured torus. In general, as discussed at the start of Section 2, we need to consider a cover of M which is a surface F homeomorphic to a sphere with three discs removed or to a torus with one disc removed, but the hyperbolic metric on F needs not have a finite volume. Thus, the once-punctured torus may be replaced by a surface with no cusp, and a closed geodesic as the boundary of its convex core, and the three-punctured sphere may be replaced by a surface with less than three cusps and some closed geodesics as the boundary of the convex core. The crucial step which allows one to proceed in the same way as in the preceding sections is to choose the cutting geodesics in F to be orthogonal to any closed geodesic boundary components of the convex core of F. Again this yields a tiling of $\mathbb{H}^2$ by quadrilaterals, but these quadrilaterals may be ultra-ideal, i.e., have vertices beyond infinity. The diagonals of an ultra-ideal quadrilateral join opposite ends and are orthogonal to any closed geodesic boundary components of the convex core of the quadrilateral.

Recall that the hyperbolic three-punctured sphere is the double of an ideal triangle. Similarly, a hyperbolic three-holed sphere F is the double of a triangle, some of whose vertices are ultra-ideal. We choose two of the common edges of these triangles to be the cutting geodesics for F, so they cut F into a quadrilateral Q with some ultra-ideal vertices which admits a reflectional symmetry in a diagonal. In particular, the diagonals for Q meet at right angles. The convex core of a one-holed torus may not admit any reflectional symmetries.

Now all the lemmas in Sections 2, 3 and 4 can be proved in essentially the same way, but the references to cusps of the quadrilaterals will need to be replaced by references to the vertices of the quadrilaterals. References to the cusp edges, which are geodesics with one end at the cusp, will need to be replaced by references to edges which go to the same vertex of the tiling of the universal covering of F. In the case of a three-holed sphere, the symmetries of the quadrilaterals show that the set of edges of the tiling with a common vertex is invariant under reflections in any of them. In the case of a one-holed torus, it has a rotational symmetry of order two. Thus, as discussed at the start of Section 4, the cutting geodesics for the tiling of $\mathbb{H}^2$ by quadrilaterals admit some limited symmetries. Namely, if one considers the family $\ldots, E_{-2}, E_{-1}, E_0, E_1, E_2, \ldots$ of those geodesics that end at a vertex of the tiling, the even-numbered ones are invariant under reflection in any of them, and the same holds for the odd-numbered ones. Further, the entire family of E_i's is invariant under reflection in the bisector of two consecutive E_i's. Finally, recall that some of our earlier arguments depended on the fact that the element α of $\pi_1(F)$ carried by the closed geodesic γ is not a parabolic element. We will also need the fact that γ cannot be a component of the convex core of F. For then γ would be simple, so that two distinct geodesics in $\mathbb{H}^2$ which lie above γ cannot cross. ∎

We can now deduce the following bounds on the self-intersection angles of a hyperbolic geodesic in terms of its length.

Corollary 5.2. *If γ is a closed oriented geodesic on an orientable hyperbolic surface, and ϕ is the angle formed by two outgoing arcs of γ at a self-intersection point, then* $\Pi(\frac{l(\gamma)}{2}) < \phi < \pi - \Pi(\frac{l(\gamma)}{4})$.

Proof. The first inequality follows immediately from Theorem 5.1. The second inequality follows by applying Lemma 2.5 to the supplementary angle $\pi - \phi$. ∎

We can also deduce the following bounds on the lengths of polygons formed by the lines above γ in $\mathbb{H}^2$.

Corollary 5.3. *If γ is a closed geodesic on an orientable hyperbolic surface M, then the triangles formed by the geodesic lines above γ in $\mathbb{H}^2$ have sides shorter than $l(\gamma)$, and the n-gons have sides shorter than $(n-2)l(\gamma)$.*

Proof. The length of a side s of a closed polygon P in $\mathbb{H}^2$ is bounded above by the sum of the lengths of the orthogonal projections of the other sides to the line containing s. These lengths are bounded by the lengths of the projections of the lines containing them to s, and in the case of the two sides adjacent to s, by half of that length. ∎

If the polygon P has an oriented boundary (with the orientations of the sides induced by an orientation of γ), then Lemmas 2.3 and 2.5 show that the sides of P are shorter than $\frac{n-1}{2}l(\gamma)$.

Now we can deduce the following result.

Corollary 5.4. *Let γ and δ be closed geodesics on an orientable hyperbolic surface M, and let l and m be distinct geodesics in $\mathbb{H}^2$ above γ and δ. Then the orthogonal projection of l onto m is shorter than $l(\gamma) + l(\delta)$.*

Interestingly, the case of two intersecting lines above distinct geodesics reduces to the case of two disjoint lines above a single geodesic, while the case of two disjoint lines above distinct geodesics reduces to the case of two intersecting lines above a single geodesic. First, we consider the situation where l and m are disjoint.

Lemma 5.5. *Let γ and δ be closed geodesics on an orientable hyperbolic surface M, and let l and m be disjoint geodesics in $\mathbb{H}^2$ above γ and δ, respectively. Then the orthogonal projection of l onto m is shorter than $l(\gamma) + l(\delta)$.*

Proof. Without loss of generality, we can assume that $l(\gamma) \leq l(\delta)$. As usual, we let α denote the element of $\pi_1(F)$ represented by γ, so that α acts on $\mathbb{H}^2$ with l as its axis. If m and αm are disjoint, the proof of Lemma 2.1 shows that the orthogonal projection of m onto l has length strictly less than $l(\gamma)$, thus proving the lemma in this case. Now suppose that m and αm cross at a point r'. Let p and p' be the closest points of l and m, so the arc $[p, p']$ is perpendicular to l and to m. Let m_∞ be the point at infinity of m such that r' lies in the ray (p', m_∞), let q be the foot of the perpendicular to l from m_∞, let q' be the intersection of qm_∞ with αm, and let s be the foot of the perpendicular to αm from m_∞. See Figure 16. To prove the lemma we need to show that the arc $[p, q]$, which is half of the orthogonal projection of m to l, is shorter than half of $l(\gamma) + l(\delta)$.

If q lies between p and αp, then $l(p, q) < l(p, \alpha p) = l(\gamma) \leq \frac{1}{2}(l(\gamma) + l(\delta))$, as required. Next suppose that q lies beyond αp, as in Figure 16. Then the angle $m_\infty q' \alpha m_\infty$ must be acute because it is equal to an interior angle of the geodesic quadrilateral $q' q \alpha p \alpha p'$ whose other interior angles are right angles. It follows that s must lie on the ray $[q', \alpha m_\infty)$,

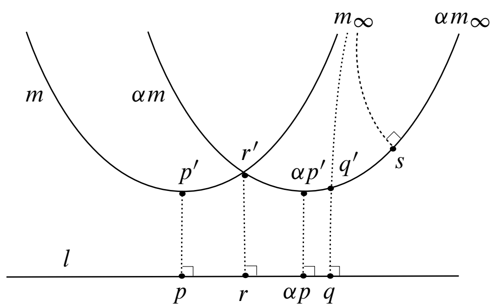


Figure 16. Orthogonal projection of geodesics above different curves.

as shown in Figure 16. Therefore, $l(r,q) < l(r',q') < l(r',s)$. As (r',s) is half of the orthogonal projection of m to αm, Theorem 5.1 tells us that $l(r',s) < \frac{1}{2}l(\delta)$. Thus $l(r,q) < \frac{1}{2}l(\delta)$. As $l(p,r) = \frac{1}{2}l(p,\alpha p) = \frac{1}{2}l(\gamma)$, it follows that $l(p,q) = l(p,r) + l(r,q) < \frac{1}{2}l(\gamma) + \frac{1}{2}l(\delta)$, as required.

If $q = \alpha p$ then $q' = \alpha p' = s$, so $l(p,q) = l(p,r) + l(r,\alpha p) < l(p,r) + l(r',\alpha p') = l(p,r) + l(r',s)$. As in the preceding case, $l(r',s) < \frac{1}{2}l(\delta)$ by Theorem 5.1, and $l(p,r) = \frac{1}{2}l(\gamma)$. Again it follows that $l(p,q) < \frac{1}{2}l(\gamma) + \frac{1}{2}l(\delta)$, as required. ∎

We now consider the case when l and m cross. In this case, we get a better bound for the projection length, which depends only on Lemma 2.1.

Lemma 5.6. *Let γ and δ be closed geodesics on an orientable hyperbolic surface M, and let l and m be distinct crossing geodesics in $\mathbb{H}^2$ above γ and δ, respectively. Then the orthogonal projection of a bisector of l and m onto m is shorter than $l(\gamma) + l(\delta)$.*

Proof. As l and m cross, so do γ and δ. The idea of our proof is to perform cut and paste on γ and δ so as to obtain a new curve of length $l(\gamma) + l(\delta)$. The shortest closed geodesic σ in the homotopy class of this curve must have $l(\sigma) < l(\gamma) + l(\delta)$.

Consider the crossing geodesics l and m in $\mathbb{H}^2$, and let p denote the intersection point $l \cap m$. A cut and paste at p determines a cut and paste operation on γ and δ at a single point, which must yield a single piecewise geodesic closed curve η with two corners at the cut and paste point. Thus, there are two piecewise geodesic paths n and n' in $\mathbb{H}^2$ above η, which pass through p. See Figure 17. Each proceeds along l from p for a distance equal to $l(\gamma)$, then turns a corner onto a translate of m, and proceeds a distance $l(\delta)$, etc. We let σ denote the closed geodesic in the homotopy class of η. Corresponding to n, we can construct a geodesic o in $\mathbb{H}^2$ above σ by simply joining the midpoints of the geodesic segments of n by geodesic segments. And similarly, we can construct a geodesic o' in $\mathbb{H}^2$ above σ by simply joining the midpoints of the geodesic segments of n'. The reason why o and o' are geodesic rather than just piecewise geodesic is that n, and hence o, is invariant under rotation through π about each of the points where n meets o, and a similar statement holds for n' and o'. As n and o have the same stabilizer, and n' and o' similarly, it follows that o and o' must lie above the closed geodesic σ. As o' is a translate of o by the action of the stabilizer of l, it follows that o and o' are disjoint. As a rotation through π about p sends o to o', their bisector b goes through the point p.

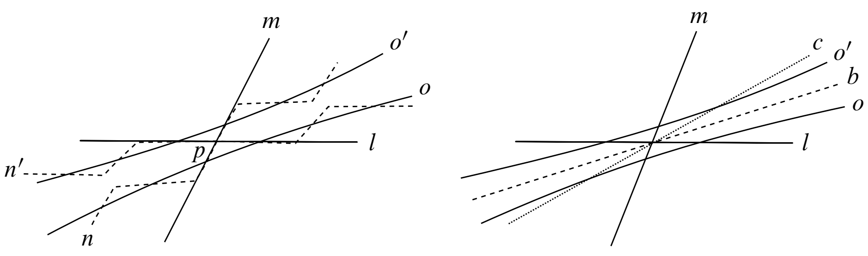


Figure 17. The geodesics o and o' arising from piecewise geodesics n and n'.

Now Lemma 2.1 says that the orthogonal projections of b to o and o' are not larger than $l(\sigma)$. These have the same length as the projections of o and o' to b, which are the same interval of b. This interval contains the feet of the perpendiculars to b from the endpoints of l, so the projection of l to b is shorter than the projections of o and o' to b. For the same reason, the projection of m to b is shorter than the projections of o and o' to b. Thus, the orthogonal projections of b to l and m are shorter than $l(\sigma)$.

Let c denote the bisector of l and m. Then one of l and m forms a larger angle with c than with b, so that the orthogonal projection of c to one of l and m is shorter than the orthogonal projection of b. Hence the orthogonal projections of c to l and m, which are equal, must be shorter than $l(\sigma)$, which is smaller than $l(\gamma) + l(\delta)$, as required. ∎

Lemma 5.6 gives a lower bound for the intersection angles of two hyperbolic geodesics.

Corollary 5.7. *The intersection angles of two closed geodesics γ and δ in an orientable hyperbolic surface are larger than* $2\Pi(\frac{l(\gamma)+l(\delta)}{2})$.

Theorem 5.4 also shows that the sides of the n-gons formed by the lines above a family of geodesics $\gamma_1, \gamma_2, \ldots, \gamma_k$ are shorter than $2(n-2)\max\{l(\gamma_i)\}$.

The bound for the lengths of the orthogonal projections given in Theorem 5.1 is optimal in the following sense:

Claim 5.8. *For each hyperbolic surface M with $\chi(M) < 0$, there is a sequence of closed geodesics γ_n and lines l_n, m_n in $\widetilde{F}$ above γ_n such that the length of the orthogonal projection of m_n to l_n, divided by the length of γ_n, converges to* 1.

Proof. As each hyperbolic surface M with $\chi(M) < 0$ has a (usually infinite) covering which is a hyperbolic sphere with three holes, it suffices to prove the claim when M is a hyperbolic sphere with three holes, each corresponding to a cusp or a boundary curve of its convex core. We can cut M along two infinite geodesics meeting the boundary curves of the convex core orthogonally to get an ideal or ultra-ideal quadrilateral. In each case, this quadrilateral has a reflectional symmetry that shows that the diagonals intersect at right angles, and the distances between opposite sides of the quadrilateral are equal. Let x and y be generators of $\pi_1(M)$ dual to the cutting geodesics. To get a sequence γ_n where l_n intersects m_n, let γ_n be the geodesic represented by the cyclic word $(xy)^n x$. Let l_n and m_n be two pre-images of γ_i in the universal cover $\widetilde{M}$ of M as in Figure 18. Since the infinite words corresponding to l_n and m_n overlap in a word $(xy)^{n-1}x$, then l_n and m_n cross together $2(n-1)$ quadrilaterals.

We claim that the arc a_n where l_n crosses these quadrilaterals is contained in the projection of m_n to l_n. To see this, consider the shaded quadrangle in Figure 18, formed by l_n, m_n, the first cutting geodesic E that they cross together and the last diagonal D that they cross together. Recall that M is the double of an ideal or ultra-ideal triangle, and so it admits a reflection symmetry which interchanges these triangles. It follows that in the universal cover $\widetilde{M}$ of M, reflection in any quadrilateral edge or diagonal is a symmetry

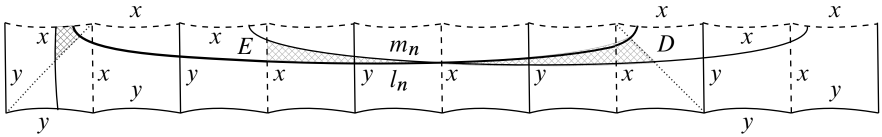


Figure 18. Two lines with long projections.

of the tiling of $\widetilde{M}$ by quadrilaterals. In particular, a reflection in D preserves the tiling and preserves but inverts the geodesic l, because it inverts the infinite word which is the unwrapping of $(xy)^n x$. It follows that the angle at the point where l_n crosses D is $\pi/2$. Also, the reflection of l_n in E is a geodesic that crosses two of the $2(n-1)$ quadrilaterals crossed by l_n and m_n and "turns north" at the third quadrilateral. So the reflection of l_n lies "north" of l_n in these quadrilaterals, and therefore the shaded angle at the intersection of l_n and E is greater than its supplementary angle, and so it is greater than $\pi/2$.

Now, if d is the distance between opposite sides of the quadrilaterals, then γ_n is shorter than $(2n+1)d$, as it is homotopic to a polygonal curve that runs through the middle arcs of the quadrilaterals, made of $2n$ arcs of length d and 2 subarcs of length $d/2$. On the other hand, the orthogonal projection of m_n to l_n is longer than the arc a_n, which has length at least $2(n-1)d$. So the ratio between the length of the projection and the length of the geodesic is larger than $\frac{2n-2}{2n+1}$, so it converges to 1.

To get a sequence γ_n where l_n is disjoint from m_n, let γ_n be the geodesic represented by the word $x(xy)^n y$, and let l_n and m_n be two pre-images of γ_i in $\widetilde{F}$, so that the infinite words corresponding to l_n and m_n overlap in the word $(xy)^{n-1}$. So l_n and m_n cross together $2(n-1)-1$ quadrilaterals. One can show as before that the arc where l_n crosses these quadrilaterals is contained in the projection of m_n to l_n. So γ_n is shorter than $2(n+1)d$, while the orthogonal projection of m_n to m_n is longer than $(2n-3)d$. Therefore, the ratio between the length of the projection and the length of γ_n is larger than $\frac{2n-3}{2n+2}$, and so it converges to 1. ∎

The upper bound for the lengths of the orthogonal projections of lines above two different geodesics given in Corollary 5.4 is also optimal. Consider the sequence of geodesics γ_n and δ_n represented by the words $(xy)^n x$ and $(xy)^{n+2} x$ in a three-holed sphere. Take two geodesic lines l_n and m_n above γ_n and δ_n whose axes have a common subword $(xy)^n x(xy)^n$. Then the projection of m_n to l_n has length greater than $4nd$, while the lengths of γ_n and δ_n are smaller than $(2n+1)d$ and $(2n+5)d$, respectively, so the ratio between the length of the projection and the sum of the lengths of γ_n and δ_n is larger than $\frac{4n+1}{4n+6}$, and so it converges to 1.

Remark 5.9. Although the bound for the projection length given in Theorem 5.1 is optimal in relative terms, it may not be so in absolute terms. The above examples suggest that the projection length for two intersecting geodesics above γ might be shorter than $l(\gamma)-c$ for some constant $c>0$ independent of γ. Comparing the angles of parallelism of $\frac{l}{2}$ and

$\frac{l-c}{2}$, one can see that $\Pi(\frac{l-c}{2})/\Pi(\frac{l}{2}) = \arctan(e^{-l/2+c/2})/\arctan(e^{-l/2})$ is an increasing function of l which is greater than 1 for each $c > 0$. For $c \approx 1.45$, this function is already greater than 2 when $l = 4\log(1+\sqrt{2}) \approx 3.5255$ (the length of the shortest non-simple geodesic on any hyperbolic surface, or twice the width of a regular ideal quadrilateral). So a gap of around 1.45 between the projection length and the geodesic length of γ would imply that the self-intersection angles of γ are larger than $2\Pi(l(\gamma)/2)$, or twice the bound given in Corollary 5.2.

Acknowledgments. We thank the anonymous referees for their helpful suggestions and corrections.

Funding. Research partially supported by PAPIIT-UNAM grant IN115414.

References

[1] A. Basmajian, The stable neighborhood theorem and lengths of closed geodesics. *Proc. Amer. Math. Soc.* **119** (1993), no. 1, 217–224 Zbl 0779.30030 MR 1152271

[2] A. F. Beardon, *The geometry of discrete groups*. Grad. Texts in Math. 91, Springer, New York, 1983 Zbl 0528.30001 MR 698777

[3] J. Gilman, A geometric approach to the hyperbolic Jørgensen inequality. *Bull. Amer. Math. Soc. (N.S.)* **16** (1987), no. 1, 91–92 Zbl 0655.30030 MR 866021

Received 26 May 2019.

Max Neumann-Coto
Instituto de Matemáticas, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Mexico; max.neumann@im.unam.mx

Peter Scott
Mathematics Department, University of Michigan, Ann Arbor, 48109 Michigan, USA; pscott@umich.edu