

THE DISTRIBUTION OF SIMPLE CLOSED GEODESICS ON A RIEMANN SURFACE

PETER BUSER AND HUGO PARLIER

ABSTRACT. The subject of this lecture is ongoing research about the distribution of the simple closed geodesics on a compact Riemann surface, endowed with the Poincaré metric of constant curvature -1 . It is well known that the closed geodesics on such a surface are dense, they are even dense in the unit tangent bundle. The situation is quite different, however, if one restricts one's consideration to the simple geodesics: Birman and Series [2] have shown that the union of all closed geodesics without self-intersections is a nowhere dense set. Our goal is to give a quantitative version of this result.

1. INTRODUCTION

It is a well known result that on a compact negatively curved Riemannian manifold M the closed geodesics are dense. They are even dense in the unit tangent bundle. The “responsible” phenomenon is that if you consider any point $p \in M$ and any unit tangent vector v at p , then a geodesic ray with p and v as initial conditions will return arbitrarily close to p and with tangent vector arbitrarily close to v , if you only wait long enough. It is then possible to slightly modify p and v such that at the moment of return the initial and end conditions match exactly and one has a smooth closed geodesic.

Now let us assume that M is a compact *Riemann surface* of genus $g \geq 2$, i.e. a compact two dimensional orientable Riemannian manifold of genus g endowed with a Riemannian metric of constant curvature -1 . We shall also say that M is a *hyperbolic surface*.

In this case, one may well imagine that if we take $p \in M$ at random and let a geodesic ray γ start at p into an arbitrary direction v , then it will be difficult for γ to come back near p over and over again without ever crossing itself. Hence, we expect that most of the closed geodesics on M have self-intersections (i.e. transversal crossings). This is indeed true and the following results are known.

Theorem A (Mirzakhani). *Let $\mathcal{N}_M(s, L)$ be the number of simple closed geodesics of lengths $\leq L$ on M . For $L \rightarrow \infty$ this number has the asymptotic behavior*

$$\mathcal{N}_M(s, L) \sim c_M L^{6g-6},$$

where c_M is a constant depending on M .

This theorem—which first was shown by McShane and Rivin [6, 7] in the case of the once-punctured torus and for which there is also a weaker form in [9]—is in sharp contrast

1991 *Mathematics Subject Classification*. Primary 30F45, 53C22, Secondary 30F45, 30F60.
Key words and phrases. Riemann surfaces, closed geodesics, moduli space.

to Huber's asymptotic law [4],

$$\mathcal{N}_M(L) \sim e^L/L,$$

where $\mathcal{N}_M(L)$ is the number of *all* closed geodesics of lengths $\leq L$ on M .

The second result, due to Birman and Series [2], says that not only are the simple closed geodesics much fewer in number, they also pass only through very special locations:

Theorem B (Birman-Series). *The set $S \subset M$ of all points on M which lie on a simple closed geodesic is nowhere dense and has Hausdorff dimension one.*

(The statement is true, more generally, for the sets S^k of points on M which lie on a complete—closed or none closed—geodesic that has at most k transversal self-intersections; $k = 0, 1, \dots$.) Hence, there is an open dense subset on M that is completely avoided by the simple closed geodesics. Our aim is to show that the avoided set has even some thickness that is independent of the geometry of M :

Theorem C. *There exists a constant $c_g > 0$, depending only on g , such that any compact Riemann surface M of genus g contains a disk of radius c_g into which the simple closed geodesics do not enter.*

As an illustration we show the distribution of closed geodesics on a one-holed torus T (a surface of genus 1 with one closed boundary geodesic). Although this is not a closed surface the phenomenon shows up as well, and the computation of the geodesics was easier to carry out on it. The figures represent the right-angled geodesic octagon \tilde{T} in the upper half plane model \mathbb{H} of the hyperbolic plane that is obtained by cutting T open along two disjoint geodesic arcs which meet the boundary of T perpendicularly at their endpoints. Figure 1 shows the paths on \tilde{T} of the first 10, respectively the first 50 closed geodesics of T . Figure 2 shows the paths of the first 10, respectively 50, simple closed geodesics

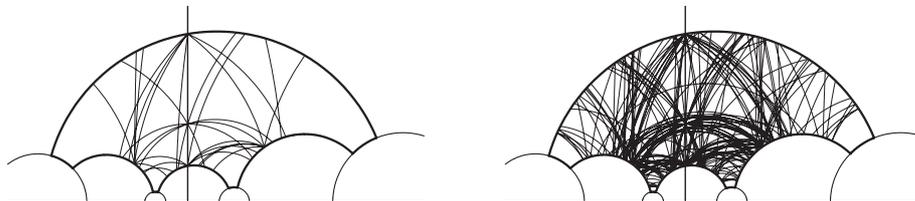


FIGURE 1. The first 10, respectively 50, geodesics of a one-holed torus drawn on a fundamental domain in the upper half plane

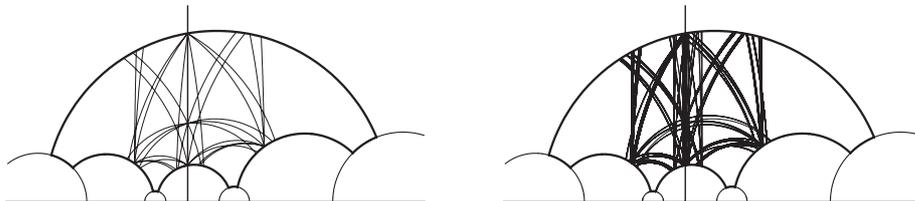


FIGURE 2. The first 10, respectively 50, simple closed geodesics

geodesics. "First n " is meant with respect to the combinatorial enumeration procedure that we used for the drawing algorithm. In both cases the full set of geodesics is still denser, but the difference in behavior is evident.

In the next paragraphs we give an outline proof of Theorem C, based on figures, in which we try to bring the geometric phenomena into evidence.

2. NEAR THE BOUNDARY OF MODULI SPACE

Example 1 (thrice punctured sphere). Let us begin with an example where the statement of Theorem C becomes quite evident. The example itself is a limit case and thus lies on the boundary of the scope of Theorem C. It is obtained as follows.

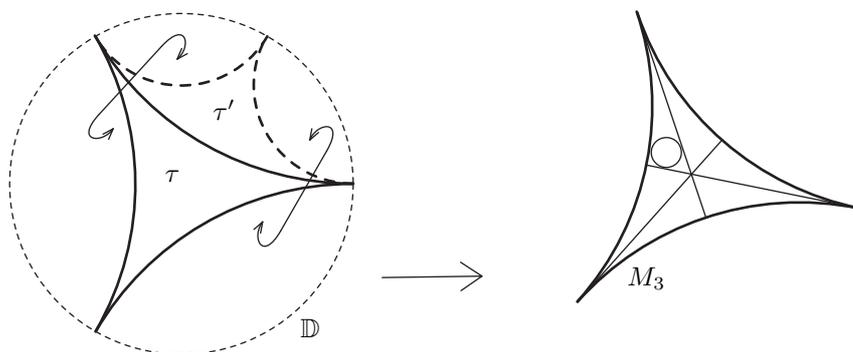


FIGURE 3. An ideal hyperbolic quadrilateral and the resulting thrice punctured sphere

Let \mathbb{D} be the unit disc model of the hyperbolic plane and let τ be an ideal geodesic triangle, i.e. a geodesic triangle domain in \mathbb{D} whose vertices lie at infinity (the boundary of \mathbb{D}) as shown on the left hand side of Fig. 3. Let Γ be the group of isometries of \mathbb{D} (Möbius transformations) generated by the three reflections across the sides of τ , and let $\Gamma_3 \subset \Gamma$ be the index two subgroup consisting of all orientation preserving elements of Γ . Then τ is a fundamental domain of Γ and the union $\tau \cup \tau'$ is a fundamental domain of Γ_3 , where τ' is the reflected image of τ across one of its sides. Our example—shown on the right-hand side of Fig. 3—is the surface $M_3 = \mathbb{D}/\Gamma_3$, the quotient of \mathbb{D} by the action of Γ_3 . The surface M_3 may also be described as being obtained by a pasting procedure; it is the surface obtained by gluing together the sides of the ideal quadrilateral $\tau \cup \tau'$ as shown on the left-hand side in Fig. 3.

M_3 does not have any simple closed geodesics. But it has six simple complete geodesics, as shown in the figure. Three of them (used as contour lines in the figure) connect the vertices at infinity with each other, the other three go from a vertex at infinity to a contour line and from there back to the same vertex at infinity. The complement of these geodesics consists of 12 triangles. The inscribed circular discs have the property that no simple complete geodesics enter into them.

We shall show in this paragraph that if a surface lies close to the boundary of moduli space, then the situation on it is slightly similar to this example.

Example 2 (flat torus). Now let us give a counterexample to Theorem C, in order to show an other extremal case. Of course, this example will not be a hyperbolic surface.

The surface here is a flat torus, F , obtained by gluing together opposite sides of a Euclidean rectangle as shown in Fig. 4. Any complete closed or non-closed geodesic on F is simple, and for any $\varepsilon > 0$ there exists on F a closed geodesic that is ε -dense, i.e. for which any point on F lies within distance less than ε .

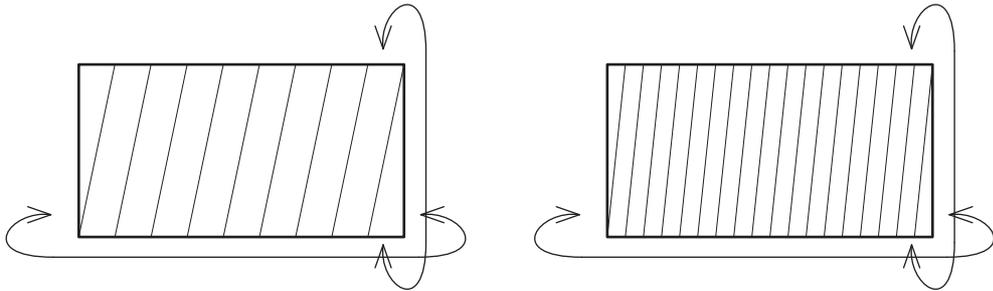


FIGURE 4. Closed geodesics on a flat torus

For a hyperbolic surface in general the situation is somewhere between these two examples. In the present section we look at the case where, to speak in terms of moduli, surface M lies near the boundary of moduli space. Boundary points of the moduli space \mathcal{M}_g of compact Riemann surfaces of genus g , $g \geq 2$, are surfaces with cusps, and so what we assume is that M has at least one sufficiently small closed geodesic.

It is well known that any $M \in \mathcal{M}_g$ admits a decomposition into $3g - 3$ three-holed spheres, so-called *pairs of pants*, along a set of pairwise disjoint simple closed geodesics $\gamma_1, \dots, \gamma_{2g-2}$. By a theorem due to Bers ([1, 3]), there even exists such a decomposition with all γ_i shorter than a constant L_g , the *Bers constant*, that depends only on g . Furthermore, if M contains simple closed geodesics of lengths less than $\text{arcsinh}(1)$, then one may choose a Bers decomposition in such a way that they all take part in it. Hence, for $M \in \mathcal{M}_g$ sufficiently close to the boundary, Theorem C is settled by the following.

Lemma 1. *There exist constants ε_L and r_L such that if $Y \subset M$ is a pair of pants with one boundary geodesic smaller than ε_L and the others smaller than L , then Y contains a disk of radius r_L into which the simple closed geodesics of M do not enter.*

Proof. (Sketch) Let us consider again a limit case: assume that Y has one cusp (instead of a boundary geodesic of length $< \varepsilon_L$) and two boundary geodesics of lengths $< L$. Figure 5 shows Y on the left-hand side, and a lift \tilde{Y} of Y in the universal covering \mathbb{D} of M on the right-hand side. \tilde{Y} is a geodesic polygon domain with four right angles and two vertices at infinity; it may also be obtained by cutting Y open along the two perpendicular geodesic arcs from the cusp to the two boundary geodesics γ_1, γ_2 .

Now consider the parts on Y of the simple complete geodesics of M . There is one example entirely on Y (not drawn in the figure) that goes from the cusp to the common perpendicular of γ_1 and γ_2 , and from there back to the cusp; its lift in \tilde{Y} is the geodesic that connects the two vertices e and f at infinity with each other. Denote by a, d ,

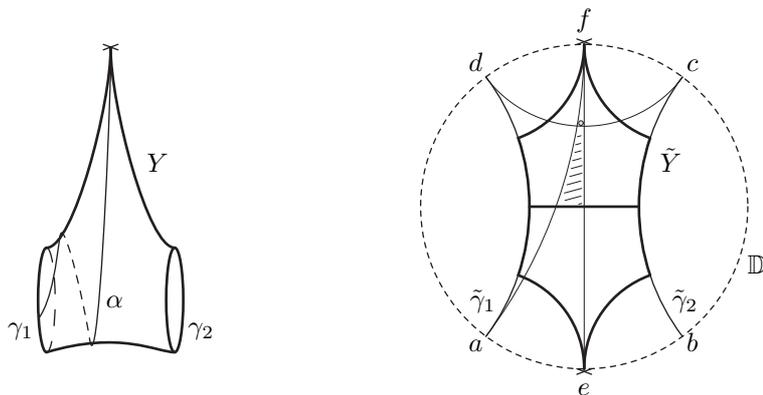


FIGURE 5. Text

respectively b, c the ideal endpoints at infinity of the lifts $\tilde{\gamma}_1, \tilde{\gamma}_2$ of γ_1, γ_2 . If α is any complete simple geodesic on M that has an ideal endpoint at the cusp of Y , and if $\tilde{\alpha}$ is a lift of α in \mathbb{D} that intersects \tilde{Y} , then $\tilde{\alpha}$ is either the geodesic from e to f , or it goes from e or f to one of the geodesics $\tilde{\gamma}_1, \tilde{\gamma}_2$. It can therefore not enter into the shaded area shown in the figure (part of triangle $ae f$).

If α on M intersects Y but does not go into the cusp, then it must connect γ_1 with γ_2 . For any lift $\tilde{\alpha}$ of it in \mathbb{D} the intersection $\tilde{\alpha} \cap \tilde{Y}$ must therefore be a subset of the ideal quadrilateral $abcd$. Now, part of our shaded area lies outside this quadrilateral! So, this part contains a disk of some positive radius into which no lift of a simple geodesic can enter.

It is not hard to see that the same argument holds true if the assumption of a cusp is replaced by the assumption that Y has a sufficiently small boundary geodesic. Hence, the lemma. \square

3. CONTINUITY OF THE RADII

For general $M \in \mathcal{M}_g$ the proof of Theorem C uses a continuity argument. Let c_M be the radius of the largest disk on M into which no simple closed geodesic enters. By the Birman-Series theorem (Theorem B), c_M is positive for each M . If we can show that c_M is a continuous function of M , $M \in \mathcal{M}_g$, then we are done because, as is well known, the subset $\mathcal{M}_{g,\varepsilon}$ of \mathcal{M}_g consisting of all surfaces with shortest closed geodesic of length $\geq \varepsilon$ is compact. Hence, c_M has a positive lower minimum in $\mathcal{M}_{g,\varepsilon}$ and, by Lemma 1, a positive lower bound outside $\mathcal{M}_{g,\varepsilon}$.

The continuity is understood with respect to the topology of \mathcal{M}_g based on the following concept of quasi-isometry (there are also variants). A one-to-one mapping $\phi : M \rightarrow M'$ is called a k -quasi-isometry ($k \geq 1$) if for any $p, q \in M$ one has

$$\frac{1}{k} \cdot \text{dist}(p, q) \leq \text{dist}(\phi(p), \phi(q)) \leq k \cdot \text{dist}(p, q).$$

The *distance* of $M, M' \in \mathcal{M}_g$ is then defined as the infimum over all $\log(k)$ for which a k -quasi-isometry from M to M' exists.

Now assume that there is a k -quasi-isometry $\phi : M \rightarrow M'$, with k close to 1, and consider a disk on M into which no simple closed geodesic enters. Its image under ϕ in M' is not exactly a disk but it contains a disk of almost the same radius. And the images, $\phi(\alpha)$, of the simple closed geodesics α of M do not enter in it. But where are the closed geodesics in the free homotopy classes of the images? Do they lie almost at the same place, as shown in Fig. 6, which means that we have found a "good" disk of

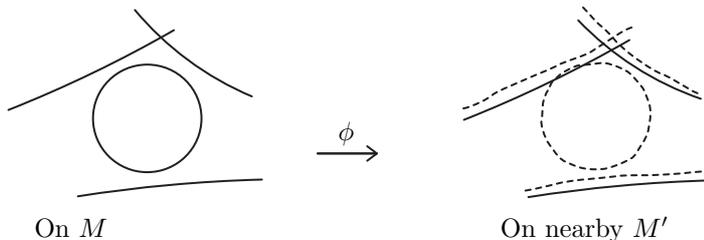


FIGURE 6

almost the same radius on M' , or is the situation rather as in the next figure, where the geodesic α' in the homotopy class of $\phi(\alpha)$ has moved too far away? The next example

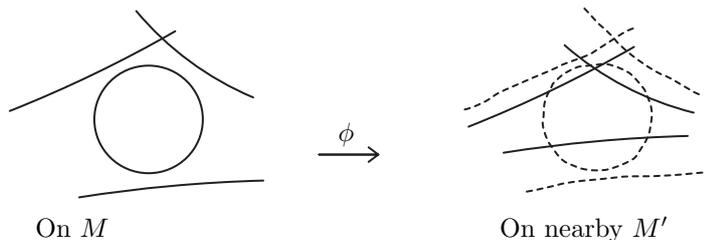


FIGURE 7

shows that there are, in fact, situations where α' slips away from $\phi(\alpha)$.

Example 3. Let S and S' be surfaces of revolution in \mathbb{R}^3 as shown in Fig. 8, where a contour line is rotated about a horizontal axis. If the contour line is convex, then the surface has negative curvature. In both cases the contour line is almost straight, and there exists a k -quasi-isometry ϕ from S to S' with k almost equal to 1 that sends boundary component γ_i of S to boundary component γ'_i of S' , $i = 1, 2$. The contour lines are such that on S as well as S' there exists a unique simple closed geodesic α , respectively α' homotopic to the boundary. This geodesic lies at the narrow part of the



FIGURE 8. Quasi-isometric surfaces of revolution

surface. On S the narrow part is near γ_1 ; on S' it is near γ'_2 . Hence, α' lies far away from $\phi(\alpha)$.

The next lemma will tell us that on a hyperbolic surface the situation is different. Its statement is in terms of quasi-geodesics. If α is a geodesic parametrized with unit speed and $\phi : M \rightarrow M'$ a k -quasi isometry, then $\phi \circ \alpha$ is a curve satisfying the following definition for $q = k^2$ (again there are variants of this definition): a parametrized curve $c : I \rightarrow R$ from an interval I to a Riemannian manifold R is called a q -quasi-geodesic if for any $s, t \in I$ with sufficiently small $|t - s|$, the following holds, where $\ell(c|_{[s,t]})$ denotes the length of the arc of c corresponding to the interval $[s, t]$:

$$\ell(c|_{[s,t]}) \leq q \cdot \text{dist}(c(s), c(t)).$$

The following lemma concludes the proof of Theorem C:

Lemma 2. *For some positive function $f :]1, \infty[\rightarrow \mathbb{R}$ with $\lim_{t \rightarrow 1} f(t) = 0$ the following holds. Any closed q -quasi geodesic γ' on a hyperbolic surface lies within distance $f(q)$ from the closed geodesic γ in its free homotopy class.*

Proof. (Sketch) As homotopies on the surface may be lifted to homotopies on the universal covering we shall prove the lemma for the case where γ is a geodesic in the hyperbolic plane with ideal endpoints, say a and d , at infinity, and γ' is a q -quasi geodesic having the same ideal endpoints. We have to estimate the distance from any point p on γ' to γ .

To this end we look at the inscribed geodesic polygon $\mathcal{P} = \cdots p_{-2}p_{-1}p_0p_1 \cdots$ of γ as shown in Fig. 9 that has point p among its vertices, say such that $p = p_0$, and all of whose sides have length 1 (we might also use some other length). Any three consecutive points

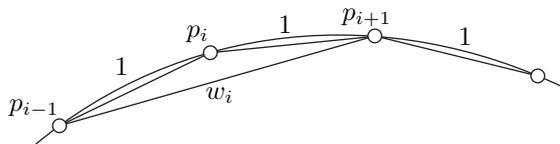


FIGURE 9. An inscribed polygon

p_{i-1}, p_i, p_{i+1} form a triangle whose longest side has length $w_i \geq 2q$ (because the length of the arc on γ' from p_{i-1} to p_{i+1} is at least 2). Now draw a comparison polygon \mathcal{P}^q as shown in Fig. 10, also with all sides of length 1, and such that any three consecutive points form a triangle whose largest side has length $2q$. The position of \mathcal{P}^q must be such that one of its sides coincides with side p_0p_1 and such that the polygon “looks towards” γ . By construction the angles at the vertices of \mathcal{P}^q are smaller than or equal to the angles of \mathcal{P} . This implies that γ separates p from the geodesic γ^q that connects the ideal endpoints b, c of \mathcal{P}^q with each other. In particular the distance from p to γ is smaller than or equal to the distance from p to γ^q . The latter is a function of q (all points of \mathcal{P}^q lie on a parallel curve of γ^1) that converges to 0 as $q \rightarrow 1$, and the lemma follows. \square

4. LARGE GENUS

In this short paragraph we want to point out that constant c_g does not have a positive lower bound independent of g .

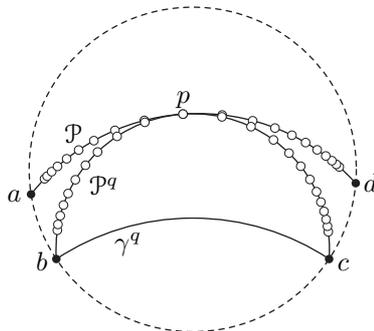


FIGURE 10. Comparison polygon and the geodesic that connects its ideal endpoints

The argument is as follows. Take any compact Riemann surface M and let $\delta > 0$. Since the closed geodesics on M are dense, there exists a sequence $\gamma_1, \dots, \gamma_n$ of them on M such that any point p lies within distance $< \delta$ of some γ_i . By a theorem of Scott [10], there exists a finite covering surface \hat{M} of some genus \hat{g} such that all primitive lifts of $\gamma_1, \dots, \gamma_n$ on \hat{M} are simple geodesics. It follows that any point on \hat{M} lies within distance $< \delta$ of some simple geodesic, and thus $c_{\hat{g}} < \delta$.

REFERENCES

- [1] Bers, L. *Finite dimensional Teichmüller spaces and generalizations*. Bull. Amer. Math. Soc., **5**(2) (1981), 131–172.
- [2] Birman, J.S. and Series, C. *Geodesics with bounded intersection number on surfaces are sparsely distributed*. Topology **24** (1985), 217–225.
- [3] Buser, P. and Seppälä, M. *Symmetric pants decompositions of Riemann surfaces*. Duke Math. J., **67** (1992), 39–55.
- [4] Huber, H. *Zur analytischen Theorie hyperbolischer Raumformen und Bewegungsgruppen*. Math. Ann. **138** (1959), 1–26.
- [5] McShane, G. *A remarkable identity for lengths of curves*. Ph.D.thesis, University of Warwick, Coventry, 1991.
- [6] McShane, G. and Rivin, I. *Simple curves on hyperbolic tori*. C. R. Acad. Sci. Paris Sr. I Math. **320** (1995), 1523–1528.
- [7] McShane, G. and Rivin, I. *A norm on homology of surfaces and counting simple geodesics*. Internat. Math. Res. Notices **2** (1995), 61–69 (electronic).
- [8] Mirzakhani, M. *Growth of the number of simple closed geodesics on hyperbolic surfaces*. to appear in Ann. of Math.
- [9] Rivin, I. *Simple curves on surfaces*. Geom. Dedicata **87** (2001), 345–360.
- [10] Scott, P. *Subgroups of surface groups are almost geometric*. J. London Math. Soc. (2) **17** (1978), 555–565.

(Peter Buser) SECTION DE MATHÉMATIQUES, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, SB-IGAT-GEOM, STATION 8, CH-1015 LAUSANNE, SWITZERLAND
E-mail address: `peter.buser@epfl.ch`

(Hugo Parlier) SECTION DE MATHÉMATIQUES, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, SB-IGAT-CTG, BCH, CH-1015 LAUSANNE, SWITZERLAND
E-mail address: `hugo.parlier@epfl.ch`