Recent Advances in Mathematics, RMS-Lecture Notes Series No. 21, 2015, pp. 45–75.

# Relations between Word Length, Hyperbolic Length and Self-Intersection Number of Curves on Surfaces

Moira Chas

Department of Mathematics, Stony Brook University, Stony Brook, NY, 11794, USA
e-mail: moira.chas@stonybrook.edu

**Abstract.** Consider an orientable surface $\Sigma$ with negative Euler characteristic, a minimal set of generators of the fundamental group of $\Sigma$, and a constant curvature $-1$ metric on $\Sigma$. Each unbased homotopy class $C$ of closed oriented curves on $S$ determines three numbers: the word length (that is, the minimal number of letters needed to express $C$ as a cyclic word in the generators and their inverses), the minimal geometric self-intersection number, and finally the geometric length. These three numbers can be explicitly computed (or approximated) using a computer.

We will discuss relations between these numbers and their statistical structure as length becomes large.

**Keywords.** Hyperbolic, surfaces, geodesics, length, word length, intersection number.

**2010 Subject Classification:** Primary 57M50

## 1. Introduction

The following can be viewed as a kind of experimental non commutative number theory which is made possible by the advent of computers. Many patterns in curves on surfaces are observed by computation. The pattern in figure 1 is one of those.

This computation lead to the proof [12] of

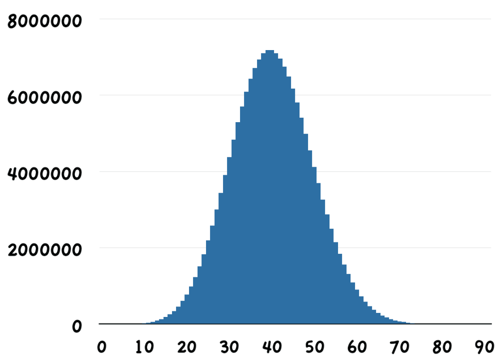


**Figure 1.** Histogram of all (about 175,000,000) non-power, free homotopy classes of word length $L = 20$ in the punctured torus, organized by self-intersection number. The mean of the self-intersection number is $400/9 \sim 45$.

**Theorem 1.1.** *On a surface with non-empty boundary and negative Euler characteristic $\chi$, the proportion of words $w$ of word length $L$ such that*

$$a < \frac{\mathrm{SI}(w) - \kappa \cdot L^2}{\sigma L^{3/2}} < b$$

*converges as $L$ goes to infinity to $\frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-\frac{x^2}{2}}\, dx$ where*

$$\kappa = \frac{\chi}{3(2\chi - 1)} \quad \textit{and} \quad \sigma^2 = \frac{2\chi(2\chi^2 - 2\chi + 1)}{45(2\chi - 1)^2(\chi - 1)}.$$

*In other words, when $L$ is very large, the distribution of self-intersection of all free homotopy classes of word length $L$ is close to a Gaussian with mean $\kappa \cdot L^2$ and standard deviation $\sigma \cdot L^{3/2}$.*

*Remark 1.2.* The expected value of intersections of $n$ random chords in a circle is $\frac{n(n-1)}{6}$ and the variance is $\frac{n(n-1)(n+3)}{45}$, [38, Chapter 6] (Here, a random chord is determined by two points independently and randomly placed on the circumference, with uniform distribution). Compare to the following: when the Euler characteristic $\chi$ is very large, the mean of self-intersections of all classes of word length $L$ is close to $\frac{L^2}{6}$ when $L$ is large and the variance is close to $\frac{L^3}{45}$.

## 2. Surfaces: Topology and geometry

### 2.1 *Topology: Surface words*

Consider an orientable surface $\Sigma$ with non-empty boundary and negative Euler characteristic. (We choose to work with surfaces with boundary to simplify the discussion, but many aspects discussed in this subsection could be repeated for closed surfaces). Choose a maximal set of disjoint arcs each starting and ending in the boundary, such that the surface minus the union of the arcs is connected (see figure 2, left). Note that this connectivity plus maximality implies that no two of the arcs are homotopic keeping endpoints in the boundary (i.e., rel boundary).

Label one side of each arc with a letter $x$, and the other side with the letter $\bar{x}$. The choice of arcs determines a minimal set of generators of the fundamental group

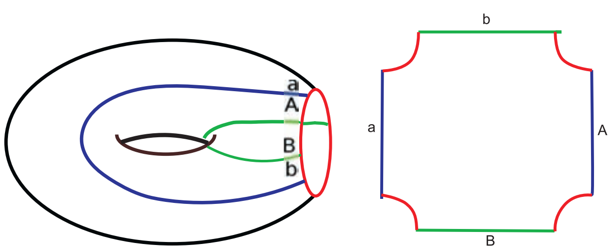


**Figure 2.** Torus "cuts"

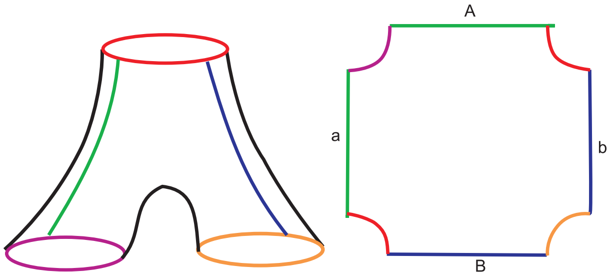


**Figure 3.** Pairs of pants "cuts"

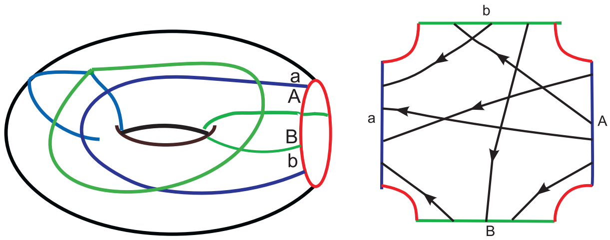


**Figure 4.** Two representatives of generators of the fundamental group of the torus with one boundary component and the corresponding labeled arcs (left) and a representative of the curve $aaabaBB$ (right)

of $\Sigma$: Choose a basepoint $P$ in the surface in the complement of the chosen arcs. A representative of the generator labeled by $x$ is a curve that starts at $P$ crosses the arc labeled $x$ from the side labeled by $x$ to the side labeled by $\bar{x}$ and goes back to $P$ without crossing any other arc. (See figure 4, left.) By cutting the surface along these arcs, a polygon is obtained with four times the number of sides as the number of generators (see figure 2). Alternating edges are labeled. By reading the labels of the edges in cyclic order a cyclic word is obtained (*cyclic* means up to cyclic permutation). This is called the *surface word*. In the example of figure 2 the surface word (written linearly) is $abAB$. In the example of figure 3 the surface word (written linearly) is $aAbB$.

It is not hard to see that the Euler characteristic of the surface is $1 - n$ where $n$ is the number of generators of the fundamental group, or the number of chosen arcs.

Our example of a torus with one boundary component and the pair of pants are exceptional in the sense that there is (up to obvious isomorphisms) only one surface word (constructed as above) for each of these surfaces. In general, more than one surface word can yield the same surface. For instance, $abABcdCD$ and $abcdABCD$ both yield a genus two surface with one boundary component.

Up to this point, we have not used the fact that one edge is labeled on one side with $x$ and the other side with $\bar{x}$. This will come into play when we discuss curves on surfaces. We will see then that the surface word not only encodes the topological information of the surface (genus and number of boundary components) but also encodes implicitly the structure of intersection of curves on the surface.

### 2.2 *Geometry: Hyperbolic metrics on surfaces*

A *hyperbolic metric* on a surface is a metric of constant curvature $-1$. By a *hyperbolic surface* we mean a surface with negative Euler characteristic, with a complete (in the sense that every Cauchy sequence has a limit) hyperbolic metric, such that if the surface has non-empty boundary, then all boundary components are geodesics (see Section 3.3 for a definition of geodesic.) We usually assume the surface is compact but some statements below allow finitely many punctures of such hyperbolic surface. In this case the hyperbolic metric has to be increased to restore completeness. Each puncture produces an infinitely long cusp like end.

If the polygon in figure 2 is drawn in the hyperbolic plane with right angles and with the natural length conditions one can glue up the sides to obtain a torus with one smooth geodesic boundary, (see also Appendix A). In a similar way, closed surfaces and other surfaces with boundary with a hyperbolic metric can be obtained, for more details, see [7, Chapter 3].

## 3. Growth of the number of closed curves on surfaces: Topology and geometry

### 3.1 *Free homotopy classes of closed curves on surfaces*

We are interested in studying equivalence classes of closed directed curves on the surface $\Sigma$ up to continuous deformation. Consider two closed oriented curves $a$ and $b$ on $\Sigma$, that is, two maps $a$ and $b$ from the oriented circle to $\Sigma$. The curves $a$ and $b$ are said to be *freely homotopic* if there exists a map from a cylinder $C$ to $\Sigma$ such that the restriction of this map to one of the (oriented) boundary components of $C$ coincides with $a$ and the restriction to the other, coincides with $b$.

The set of equivalence classes under this relation is the set of *free homotopy classes of closed curves on* $\Sigma$ and will be denoted by $\pi_0$.

There is a natural bijection between $\pi_0$ and the set of components of the space of maps from the circle to $\Sigma$, with the compact-open topology. This is the reason why the set of free homotopy classes is denoted by $\pi_0$. (This bijection holds in spaces more general than surfaces, namely, path connected spaces).

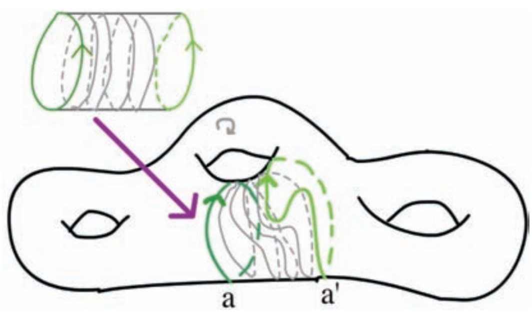


**Figure 5.** A (free) homotopy between the curves $a$ and $a'$

Here is another interpretation of $\pi_0$, (see [10]), which holds for path-connected spaces.

**Proposition 3.1.** *If $\Sigma$ is a connected surface then there is a bijection between the set of free homotopy classes of closed directed curves on $\Sigma$ and the set of conjugacy classes of $\pi_1(\Sigma, x_0)$.*

## 3.2 *Topology: Growth of the number of free homotopy classes of curves by word length*

Consider a representative of a free homotopy class of curves that intersects the union of the arcs in the smallest possible number of points. (Intersections are counted with multiplicity). The free homotopy class is labeled by a cyclic reduced word obtained by recording the arcs (and sides) the curve crosses as one traverses the directed curve. (*Cyclic* means that words are considered up to cyclic permutation, *reduced* means that no letter $x$ and its inverse $\bar{x}$ appear consecutively in the word, or any cyclic permutation of the word). In figure 6 an example of a cyclic reduced word is exhibited.

The class of the curve in figure 4, (right) is labeled by the cyclic word $aaabaBB$. (Observe that $aaabaBB$ is a *curve word* as opposed to the *surface words* such as $abAB$ associated with the torus with one boundary component.) By Proposition 3.1, the cyclic word $aaabaBB$ labels a conjugacy class of the fundamental group of $\Sigma$. (Recall that the fundamental group of a surface with boundary is a free group.) So elements in the fundamental group can be thought of as reduced words in the generators and their inverses, while conjugacy classes can be thought of as reduced cyclic words.

Hence, once a set of generators of the fundamental group is chosen, each free homotopy class of curves can be associated with a positive integer: the number of letters in the shortest word in the conjugacy class.

The next result is not hard to prove.

**Proposition 3.2.** *The total number $N(L)$ of free homotopy classes of word length $L$ is asymptotic to $(2d-1)^L/L$, where $d$ is the number of generators of the fundamental group. Namely, $\frac{(2d-1)^L/L}{N(L)} \to 1$ as $L \to \infty$.*

*Remark 3.3.* The case of closed surfaces is more complicated, see [26] and [40] and [41].

Wroten is working on precise estimates and he reported to us that the total number of free homotopy classes in the fundamental group of a surface of genus two of word length $L$ in a set of minimal generators is asymptotic to $t^L$ where $t$ is approximately

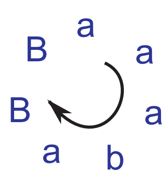


**Figure 6.** A cyclic reduced word

6.98. Note that the growth rate in the case of a surface of genus two and one boundary component is 7. This shows that the one relator for the closed surface does not affect the result too much when $L$ is relatively small.

*Remark 3.4.* In this work, we are mostly concerned with growth rate of the set of conjugacy classes of the fundamental group of a surface, (which is in natural bijection with the set of free homotopy classes of closed directed curves). There is an extensive literature about growth rate in the number of elements in a group, as a function of the word length. This corresponds to based homotopy classes. Milnor [27] proved that in compact negative curved manifolds the growth of the number of elements of the fundamental group is exponential. (The reader could check the result [18] of the always surprising Gromov.) In our setting, it follows from the work of Cannon [8] and Floyd and Plotnick [17] (see also [9, Section 7]) that the growth rate of the number of based homotopy classes of curves in a closed surface of genus two is the largest real root of the polynomial $x^4 - 6x^3 - 6x^2 - 6x + 1$, $\frac{1}{2}(\sqrt{(3+\sqrt{17})^2 - 4} + \sqrt{17} + 3)$ which is approximately, 6.98.

For compact manifolds Milnor [27] related the growth of the volume of the ball of radius $R$ to the growth rate of the fundamental group, see Theorem 3.13.

### 3.3 *Geometry: Growth of the number of geodesics by geometric length*

In this Subsection, we will assume that $\Sigma$ is an orientable surface, with or without empty boundary, with negative Euler characteristic.

Recall that a curve is a *geodesic* if it realizes the shortest path between two close-by points on the curve.

Since $\Sigma$ has negative Euler characteristic, it can be endowed with a hyperbolic metric, (recall Subsection 2.2). Hyperbolic metrics produce "optimal" representatives of free homotopy classes of curves, in the following sense:

**Theorem 3.5.** *Each free homotopy class contains a unique closed geodesic representative (unless it wraps around a puncture).*

*Remark 3.6.* Theorem 3.5 does not hold for hyperbolic metric with punctures. Indeed, there is no geodesic in the class of curves that wrap around a puncture.

See [7, Proposition 1.6.6] for a proof of Theorem 3.5. This unique geodesic representative, the *geodesic*, plays a role similar to that of a straight line on the surface: namely, if two points in the geodesic are close enough, the shortest path among all arcs joining the points is the arc of the geodesic between them.

We can ask what is the growth rate of the cardinality of the set of all closed geodesics up to geometric length $L$. Many researchers contributed to the answer of this problem among them, Huber, Margulis, Randol, Selberg. Denote by $\mathcal{C}_\Sigma(L)$ the cardinality of the set of closed geodesics on $\Sigma$, which are not proper powers of other geodesics and of length smaller than or equal to $L$. (Geodesics which are not proper powers of other geodesics are often called *primitive*.) This is a finite set [7, Theorem 1.6.11]. See [7, Theorem 9.4.14] for a proof of the following result.

**Theorem 3.7.** *Prime Number Theorem for Hyperbolic Closed Surfaces:*

$$\mathcal{C}_\Sigma(L) \sim \frac{e^L}{L}.$$

Here, $a(L) \sim b(L)$ means $\lim_{L\to\infty} \frac{a(L)}{b(L)} = 1$.

*Remark 3.8.* By setting $\ell = e^L$, Theorem 3.7 translates to $\mathcal{C}_\Sigma(\ell) \sim \frac{\ell}{\log(\ell)}$. Observe the analogy with the Prime Number Theorem, which states

$$\text{Cardinality of } \{p \text{ prime}, \ p \leq \ell\} \sim \frac{\ell}{\log(\ell)}$$

(non-power geodesics can be thought of as "prime" geodesics.)

For more on this topic, see [7, Section 9.4] and references therein.

**Theorem 3.9.** *Prime Number Theorem for Hyperbolic Surfaces with Geodesic Boundary:*

$$\mathcal{C}_\Sigma(L) \sim \frac{e^{hL}}{hL}.$$

Here, $h$ is the entropy of the geodesic flow. This entropy equals the Hausdorff dimension of the Cantor set at infinity which is the closure of the endpoints at infinity of the lifts of closed geodesics [39]. In the case of closed surfaces, $h = 1$. See, for instance, [37] for a precise definition of entropy. A proof of the Prime Number Theorem for Hyperbolic Surfaces with Boundary can be found in [24], see also [19]. The case of surfaces with punctures is more complicated.

*Remark 3.10.* For surfaces with boundary, $0 < h < 1$. The endpoints of (not necessarily closed) geodesics on a closed surface can be anywhere on the circle at infinity of the hyperbolic plane, while in the case of a surface with boundary, geodesics have endpoints in a Cantor set. Figure 7 illustrates this Cantor set: The Cantor set is what is left of the circle at infinity when one removes all arcs between endpoints of lifts of closed geodesics that contain no other endpoints of lifts of geodesics in between. (To see this observe that lifts of closed geodesics never exit a translate of the shaded octogon in a translate of the unlabled boundary in figure 7.)

*Example 1.* To illustrate these ideas, we considered a hyperbolic pair of pants (2.5, 2.6, 5) as in Definition A.4 and computed the geometric length of all geodesics shorter than 50 (see Appendix A.1). We computed the length of all geodesics up to word length 20 (20 is the maximum number of words that our computer could handle in a reasonable amount of time). By Corollary A.2, this set of geodesics contains all geodesics up to length 50 (because the constant $C$ of Corollary A.2 is $2.5 = \min\{2.5, 2.6, \frac{5}{2}\}$).

To estimate $h$ in this case, we solve numerically for $h$ the equation

$$\frac{e^{50h}}{50h} = 8,532,116.$$

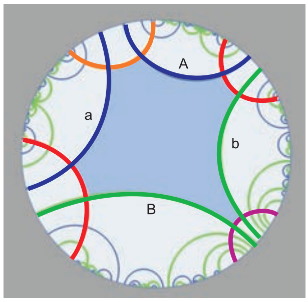


**Figure 7.** Covering space of the pair of pants (image made with Cinderella). A pair of pants is obtained by gluing pairs of sides of the shaded right angled octogon labeled with the same letter.

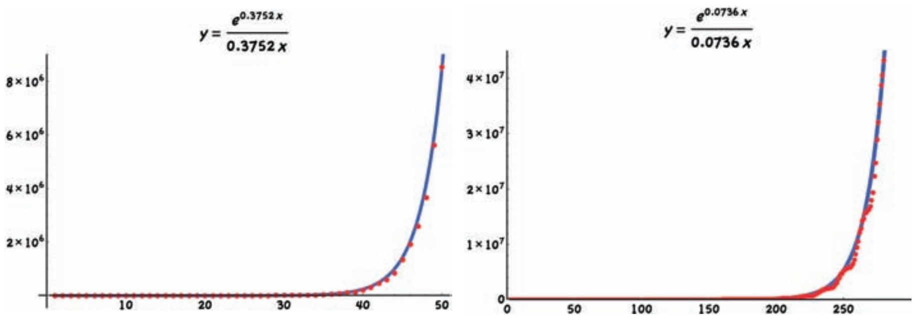


**Figure 8.** Points $(L, \mathcal{C}_\Sigma(L))$, $L \in \{1, 2, \ldots, 50\}$ and the function $F(x) = \frac{e^{0.378x}}{0.378x}$ for pair of pants (top), and $(L, \mathcal{C}_\Sigma(L))$, $L \in \{1, 2, \ldots, 280\}$ and the function $F(x) = \frac{e^{0.0736x}}{0.0736x}$ for the torus (bottom)

(8, 532, 116 is the total number of geodesics shorter than 50 in this pair of pants). We obtained 0.378 as the approximate value of $h$, which is the Hausdorff dimension of the limit set. For each integer, $L \in \{1, 2, \ldots, 50\}$, we plotted the point $(L, \mathcal{C}_\Sigma(L))$ (recall that $\mathcal{C}_\Sigma(L)$ is the total number of geodesics up to length $L$) together with the function $F(x) = \frac{e^{0.378x}}{0.378x}$ in figure 8, left.

*Remark 3.11.* In all these computations since we are only interested in counting, we consider unoriented geodesics. The number of oriented geodesics is obtained by doubling the number of unoriented geodesics.

*Example 2.* We consider the rectangular torus (14, 16) (Definition A.7) and all geodesics up to length 280, and we proceed as in Example 1. The results are displayed in figure 8, right. Here our estimation of $h$ is 0.0736. (43, 335, 144 is the total number of geodesics in this metric of length smaller than 280.) This dimension is so small because of the length of the geodesic boundary, approximately 57.227, is large (compared to the length of the generators.)

*Remark 3.12.* Observe that the growth of the number of geodesics on a closed surface is independent of the geometry and topology of the surface, while the growth of the number of geodesics on surfaces with boundary depends on the geometry of the surface.

**Table 1.** Comparison of growth of conjugacy classes and closed geodesics on surfaces with and without boundary.

| Surface | Number of closed geodesics $w$ $\mathrm{GL}(w) \leq L$ | Number of conjugacy classes $w$ $\mathrm{WL}(w) = L$ |
|---|---|---|
| Genus $g$ empty boundary | $\mathrm{e}^L/L$ | $a^L/L$, $a < (2g-1)$ $a$ is close to $2g-1$ |
| Genus $g$, $b$ boundary components | $\mathrm{e}^{hL}/hL$ | $(2g+b-1)^L/L$ |

## 3.4 *Topology and Geometry: Quasi-isometry between word metric and hyperbolic metric*

A map $f: X \longrightarrow Y$ between metric spaces $(X, d_X)$ and $(Y, d_Y)$ is a *quasi-isometry* if there exist constants $K \geq 1$ and $A > 0$ such that for each pair $x, x'$ in $X$

$$\frac{1}{K} d_X(x, x') - A \leq d_Y(f(x), f(x')) \leq K d_X(x, x') + A$$

and for all $y$ in $Y$, $d_Y(y, f(X)) \leq A$.

The constant "$A$" in the definition above is to take care of cases where the distances involved are small. When these distances are large, one can "pretend" that $A = 0$.

Fix a set of generators for the fundamental group of $\Sigma$. The fundamental group of $\Sigma$ is a metric space with distance between two words $v$ and $w$ in $\pi_1(\Sigma)$ equal to the word-length of $v \cdot w^{-1}$. Milnor-Švarc Lemma [6, Proposition 8.19] specialized in our case says the following:

**Theorem 3.13.** *Fix an element $x$ in the universal cover of $\Sigma$, $\hat{\Sigma}$. The map $q: \pi_1(\Sigma) \longrightarrow \hat{\Sigma}$ defined by $v \mapsto v \cdot x$ is a quasi-isometry ($\pi_1(\Sigma)$ has the metric given by the word length and $\hat{\Sigma}$ has the hyperbolic metric, and $v \cdot x$ is the action of the fundamental group of $\Sigma$ on the universal cover.)*

Fix a set of generators of $\pi_1(\Sigma)$ and a hyperbolic metric with geodesic boundary for $\Sigma$. For each free homotopy class $w$, write $\mathrm{WL}(w)$ for the word-length of $w$ (this is the shortest word length of a representative of the conjugacy class $w$) and $\mathrm{GL}(w)$ for the length of the geodesic in the class $w$. (Of course, these depend on the set of generators and metric, but for simplicity we do not add them to the notation). If $W \in \pi_1(\Sigma)$, the geometric length of $W$, denoted by $\mathrm{GL}(W)$ is the length of the shortest curve in $W$ passing through the basepoint of $\pi_1(\Sigma)$.

A function $f$ to the real numbers has *exponential growth* if both limits, $\limsup_{L\to\infty} \frac{\log(f(L))}{L}$ and $\liminf_{L\to\infty} \frac{\log(f(L))}{L}$ are positive and finite.

**Theorem 3.14.** *Let $K$ and $A$ be the constants given by Theorem 3.13. Then the following inclusions hold.*

$$\left\{w \in \pi_0 : \mathrm{GL}(w) \leq \frac{L-A}{K}\right\} \subset \{w \in \pi_0 : \mathrm{WL}(w) \leq L\}$$

$$\subset \{w \in \pi_0 : \mathrm{GL}(w) \leq K \cdot L + A\}$$

*Hence, if the cardinality of the set* $\{w \in \pi_0 : \mathrm{WL}(w) \leq L\}$ *grows exponentially as a function of L, that is, as* $\mathrm{e}^{a \cdot L}$ *then the cardinality of the set* $\{w \in \pi_0 : \mathrm{GL}(w) \leq K \cdot L + A\}$ *also grows exponentially possibly with a different exponential factor.*

*Proof.* Fix an element $x$ in the universal cover of the surface $\Sigma$. Consider an element in a conjugacy class $w \in \pi_0$ and a representative $W \in \pi_1(\Sigma)$ of the shortest possible word length. Recall that the distance between two elements $W$ and $V$ in $\pi_1(\Sigma)$ is $\mathrm{WL}(WV^{-1})$. In particular, the distance between $W$ and the neutral element is $\mathrm{WL}(W)$. By Theorem 3.13, there exist constants $K$ and $A$ such that

$$\frac{1}{K} d(W \cdot x, x) - A \leq \mathrm{WL}(W) \leq K d(W \cdot x, x) + A.$$

Since the projection of the geodesic segment from $W \cdot x$ to $x$ is a closed curve in the free homotopy class $w$, and the geodesic is the shortest curve in $w$, $d(W \cdot x, x) \geq \mathrm{GL}(w)$. Then

$$\frac{1}{K}\,\mathrm{GL}(w) - A \leq \mathrm{WL}(W) = \mathrm{WL}(w).$$

This implies one of the inclusions. The other can be proven similarly using the fact that, since the surface $\Sigma$ is compact, it can be assumed that the constants $K$ and $A$ of Theorem 3.13 are independent of the basepoint. Thus, one can assume that the chosen basepoint $x$ is in the geodesic. □

“conjugacy invariant” should be replaced by “quasi isometry invariant”

*Remark 3.15.* It is not hard to see that the type of growth of the number of elements in a group by word length is a quasi-isometry invariant. On the other hand, the type of growth of the number of conjugacy classes is not a conjugacy invariant [23]. In our case, we have a special situation, because the distance between the image of an element $W$ in the fundamental group and the image of neutral element in the universal cover of the surface defined in the Milnor-Svarc theorem $d(W \cdot x, x)$ is related to the length of the geodesic labeled by $W$. Thus we get $W$ on both sides of the inequality. This does not happen in general for any quasi-isometry.

*Remark 3.16.* The quasi-isometry between the fundamental group with word-length metric and the surface with the hyperbolic metric illustrates a fact that we will see repeated throughout these pages: namely, patterns that occur in the geometric realm, often have a translation in the combinatorial realm and vice versa. Theorem 3.14 is an example of this.

## 4. Self-intersection numbers of closed curves on surfaces

Each class of curves is naturally associated with a non-negative number, *the self-intersection number*. This is the smallest number of times any representative of the class crosses itself. Thus, for instance, the self-intersection number of the curves in figure 9 is 3. The *self-intersection number of a free homotopy class of curves* $w$, denoted by $\mathrm{SI}(w)$ is this minimum of all self-intersection numbers of curves in $w$. In figure 9, three representatives of the same free homotopy class are exhibited. The self-intersection number of the class is 3 (we will give an idea later why 3 is the answer).

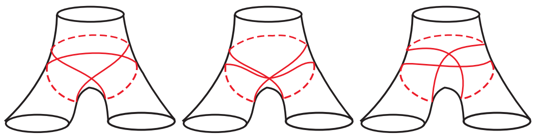

**Figure 9.** Three representatives of the free homotopy class in the pair of pants

Recall Theorem 3.5 that each free homotopy class has a unique geodesic representative.

**Theorem 4.1.** *The self-intersection number of the geodesic representative of a free homotopy class equals the self-intersection number of the class. In other words, the self-intersection number of the geodesic is the minimal number possible for any representative.*

Theorem 4.1 was noted by Poincaré in the case of self-intersection zero.

Theorem 4.1 is a consequence of [20, Theorem 2], which states that a curve $a$ that has larger self-intersection number than the minimal, has a *singular bigon* or a *singular monogon*. A *singular bigon* consists of a pair of disjoint arcs of the circle whose endpoints are mapped by $a$ to the same points and its images bound a (not necessarily embedded) disk in the surface. Now, note that if a curve has a singular bigon, then there will be two lifts of this curve to the universal cover, that bound a bigon. But in the universal cover these lifts are straight lines, and there is a unique line passing through each pair of points. Thus we get a contradiction if the geodesic representative did not have the minimal number of self-intersections.

*Remark 4.2.* The fact that a geodesic realizes the minimal intersection number shows again how intertwined are the combinatorial and geometric universes.

We see that three numbers can be associated with a free homotopy class of closed curves $w$: the self-intersection number $\mathrm{SI}(w)$, the word length $\mathrm{WL}(w)$ (provided that a set of generators of the fundamental group of the surface is chosen), and the geometric length $\mathrm{GL}(w)$ (provided that a hyperbolic metric on the surface is chosen).

*Remark 4.3.* The free homotopy class in figure 9 presents a very interesting feature: The only "picture" that can be realized by a geodesic representative is the left one. This was observed by Hass and Scott in [21], where they gave the following proof, due to Agol. If the curve on the right is the "picture" of a geodesic, then the surface minus the curve has five connected components: the three "cuffs", a triangle and a hexagon. It is not hard to see that sum of the interior angles of the hexagon equals $6\pi$ minus twice the sum of the interior angles of the triangle. Since the sum of the interior angles of the hexagon is less than $4\pi$, the sum of the interior angles of the triangle is more than $\pi$, a contradiction.

The precise notion of "picture" is given in [21]).

### 4.1 *Computing the self-intersection number from the curve words*

Birman and Series [4] found an algorithm to determine whether a free homotopy class (given as a cyclic reduced word) of curves on a surface with boundary is *simple*,

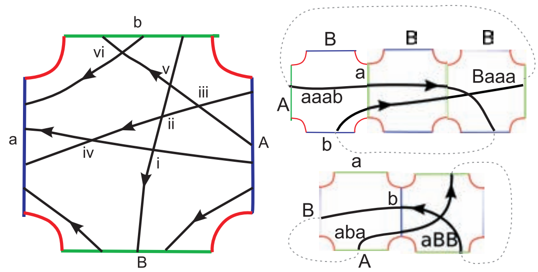


**Figure 10.** A representative of the class $aaabaBB$ in the torus with a boundary component

**Table 2.** Linked pairs of the word $aaabaBB$. The labels $i, ii, \ldots, vi$ correspond to figure 10.

| | Pair of subwords |
|---|---|
| i | *aa*, *BB* |
| ii | *aa*, *BB* |
| iii | *aab*, *Baa* |
| iv | *Baaa*, *aaab* |
| v | *aba*, *aBB* |
| vi | *aba*, *BBa* |

that is, has self-intersection number 0. Cohen and Lustig [15] extended Birman and Series method to study the intersection number of a class on a surface with boundary. Lustig [26] with analogous arguments (although more intricate) gave an algorithm to determine self-intersection numbers on a closed surface.

*Remark 4.4.* Even though the self-intersection numbers depend only on the topology of the surface, the proofs of the Cohen-Lustig-Birman-Series algorithm use hyperbolic geometry.

We will not explain the algorithm here but we give a rough idea of an equivalent form of the Birman-Series-Cohen-Lustig algorithm that appeared in [11]. The self-intersection point labeled $i$ in figure 10 is in the intersection of the arc of the curve that goes from the edge labeled by $b$ to the edge labeled by $B$, and the arc of the curve that goes from the edge labeled by $A$ to the edge labeled by $a$. The first edge corresponds to the sub-word $BB$ of the cyclic word $aaabaBB$ and the second, to the second occurrence of the sub-word $aa$ in $aaabaBB$. Figure 10, right is an illustration of another kind of pair of subwords implying an intersection.

Note that this intersection point does not depend on the "global" properties of the word. In fact, any word containing the sub-words $BB$ and $aa$ will have a self-intersection point in analogous arcs. In general, there is a one-to-one correspondence between occurrences of certain pairs of sub-words and the self-intersection points of a representative of a class (that intersects itself in the smallest possible number of points). The pairs of subwords of $aaabaBB$ corresponding to the self-intersection points of minimal representatives are listed in Table 2.

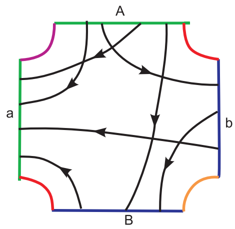


**Figure 11.** A minimal representative of the class $aaabaBB$ in the pair of pants

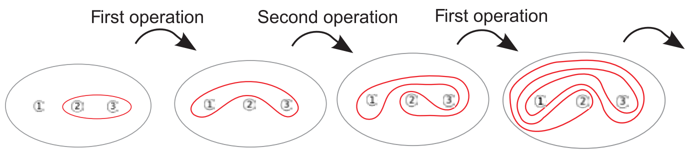


**Figure 12.** Building Thurston's simple curve

*Remark 4.5.* The pairs of sub-words that determine the self-intersection number of a curve-word depend only on the surface word. For instance, the curve word $aaabaBB$ has self-intersection 7 in the torus with one boundary component associated with the surface word $abAB$ (figure 10) and self-intersection 5 in the pair of pants associated with surface word $aAbB$.

## 4.2 *The simple curve story*

Among all closed curves on a surface, there is a subset that deserves special mention: the curves with no crossings or *simple curves*. At the end of the XIX century, 1895 Poincaré published "Analysis situs", starting topology as a new area of Mathematics [32, Translator's Introduction]. Between 1899 and 1904, he published "Compléments à l'Analysis situs" [32] correcting mistakes and tying up loose ends. In the fifth of these complements, he discussed the first characterization of simple closed curves, namely, a curve is simple if an only if all of the lifts to the universal cover are pairwise disjoint.

Despite its name, a simple curve can be far from simple (if we understand simple as opposite to complex), see for instance, figure 13. This curve was constructed by Bill Thurston in 1971. One starts with a flat disk of uncooked dough with three wooden rods standing in the dough. Label these 1, 2 and 3. Now, make a very thin closed curve with food coloring that goes tightly around rods 2 and 3. Second, perform the following two operations over and over a few times to get the curve of figure 13. The first operation is interchange rods 1 and 2 by moving them halfway around a loop in the counterclockwise direction. Then interchange 2 and 3 in a circular clockwise direction.

Here is an excerpt of a story written by Dennis Sullivan about the curve in figure 13: A couple of days later the Berkeley grad students invited me to join the painting math frescoes on the corridor wall separating their offices from the elevator foyer. While milling around before painting a grad student (Thurston) came up to ask "Do

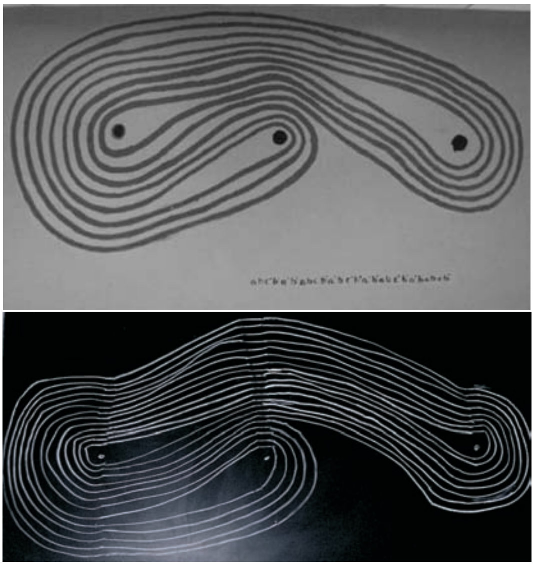

**Figure 13.** The painting on the wall the iterations of Thurston simple curve by Thurston and Sullivan (top) and a picture of further iterations of the Thurston curve by IISER Ph.D. students

you think this is interesting to paint?" It was a complicated smooth one dimensional object encircling three points in the plane. I asked "What is it?" and was astonished to hear "It is a simple closed curve". I said "You bet it's interesting!". So we proceeded to spend several hours painting this curve on the wall. It was a great learning and bonding experience. For such a curve to look good it has to be drawn in sections of short parallel slightly curved strands (like the flow boxes of a foliation) which are subsequently smoothly spliced together. When I asked how he got such curves, he said by successively applying to a given simple curve a pair of Dehn twists along intersecting curves. The "wall curve painting", two meters high and four meters wide, dated and signed, lasted on that Berkeley wall with periodic restoration for almost four decades before finally being painted over a few years ago. (See also [30]).

It is not hard to prove that the sphere with three boundary components has only three classes of undirected simple closed curves (the curves parallel to the boundary components).

All the free homotopy classes (up to direction) of simple curves on the torus with one boundary component and word length 12 are $aaaaaaaaaaab$, $aaaaaaaaaaaB$, $aababaababab$, $aaBaBaaBaBaB$, $abababbababb$, $abbbbbbbbbbb$, $aBaBaBBaBaBB$, $aBBBBBBBBBBB$. In this surface, a pattern can be guessed (and proved, see [16]). In other surfaces, the words are much more complicated. In the surface of genus two with one boundary component, associated with the surface word $abABcdCD$, there are 8362 free homotopy classes of self-intersection 0 and word length 12. Here are eight of these classes of simple curves. $abADCBcdcdCB$, $abADCBDCDCCB$, $abADCCAccdaB$, $abADCCAccdCB$, $abADCCaBaBaB$, $abAcaBABabAC$, $ababDCBabDCB$, $ababADabbcDD$.

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 5 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 2 | 10 | 8 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 6 | 17 | 14 | 13 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 2 | 12 | 28 | 34 | 38 | 22 | 8 | 8 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 6 | 20 | 46 | 76 | 79 | 50 | 52 | 40 | 20 | 12 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 2 | 12 | 32 | 84 | 140 | 146 | 140 | 158 | 134 | 98 | 66 | 40 | 24 | 6 | 6 | 2 | 2 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 6 | 20 | 52 | 146 | 226 | 264 | 330 | 371 | 358 | 339 | 274 | 206 | 120 | 91 | 72 | 33 | 22 | 6 | 4 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 2 | 12 | 32 | 92 | 232 | 354 | 480 | 638 | 782 | 866 | 878 | 856 | 694 | 538 | 522 | 388 | 264 | 188 | 100 | 62 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 20 | 52 | 156 | 345 | 550 | 818 | 1116 | 1502 | 1818 | 1998 | 2080 | 1944 | 1888 | 1786 | 1482 | 1314 | 1032 | 716 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 12 | 32 | 92 | 244 | 506 | 852 | 1298 | 1888 | 2670 | 3374 | 4042 | 4458 | 4716 | 5078 | 4966 | 4768 | 4620 | 3974 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 20 | 52 | 156 | 360 | 737 | 1282 | 1989 | 3056 | 4443 | 5836 | 7369 | 8750 | 10054 | 11208 | 12055 | 12744 | 12822 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 12 | 32 | 92 | 244 | 524 | 1076 | 1870 | 3010 | 4738 | 7002 | 9602 | 12542 | 15784 | 19132 | 22398 | 25930 | 28954 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 20 | 52 | 156 | 360 | 758 | 1548 | 2686 | 4449 | 7092 | 10664 | 15032 | 20370 | 26594 | 33450 | 41478 | 50056 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 12 | 32 | 92 | 244 | 524 | 1100 | 2184 | 3840 | 6426 | 10328 | 15798 | 22666 | 31656 | 42538 | 55416 | 71432 |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 20 | 52 | 156 | 360 | 758 | 1576 | 3050 | 5418 | 9116 | 14754 | 22761 | 33298 | 47451 | 65244 | 87852 |
| 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 12 | 32 | 92 | 244 | 524 | 1100 | 2216 | 4256 | 7550 | 12724 | 20718 | 32110 | 47774 | 69124 | 97002 |
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 20 | 52 | 156 | 360 | 758 | 1576 | 3086 | 5891 | 10412 | 17544 | 28578 | 44654 | 67144 | 98324 |

**Figure 14.** In row $L$, column $K$, we have $\mathrm{P}(K, L)$ the number of (undirected) free homotopy classes of closed curves in the **pair of pants** of word length $L$ and self-intersection $K$

In the surface of genus two with one boundary component, associated to the surface word $abcdABCD$, there are 9112 free homotopy classes of self-intersection 0 and word length 12. Here are eight of these classes of simple curves. $abadCbbDbbCd$, $abadCdaBdaBd$, $abadCdBadCbb$, $abAbaBAbCbAB$, $abAbaBcDcDcB$, $abADcDABadCd$, $abADcabADcaB$, $abAbAbaBdCdB$.

Clearly, here, patterns are much harder to find than in the punctured torus case.

## 5. Relations between word-length and self-intersection number

The Birman-Series-Cohen-Lustig, or the equivalent form [11] algorithm can be programmed. We did so and found the following tables, organizing cyclic words by word length and self-intersection. More precisely, denote by $\mathrm{P}(K, L)$ the number of undirected, non-power free homotopy classes of curves in the pair of pants of word length $L$ and self-intersection number $K$. (Note that there are exactly twice as many directed curves as undirected).

For instance, in the pair of pants with surface word $aAbB$, there are exactly two undirected, non-power, free homotopy classes of word length one and self-intersection zero, namely, $a, b$. Thus, $\mathrm{P}(0, 1) = 2$. Similarly, there is only one undirected, non-power free homotopy class of word length 2 and self-intersection 1, the "figure eight", $aB$. So $\mathrm{P}(1, 2) = 1$.

These two tables, Tables 14 and 15 exhibit many patterns. In this Subsection, as well as in Subsection 8.1 we will discuss some of the patterns we proved jointly with other authors, and point out the ones we see and are still unproven.

### 5.1 *Maximal self-intersection for a given word length*

The next result is proven in [14, Theorem 1.7].

**Theorem 5.1.** *For each free homotopy class of curves $w$ in the pair of pants,* $\mathrm{SI}(w) \geq \left\lfloor \frac{\mathrm{WL}(w)-2}{2} \right\rfloor$, *where $\lfloor x \rfloor$ denotes the largest integer smaller than or equal to $x$. This bound is sharp.*

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 5 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 8 | 4 | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 4 | 8 | 16 | 20 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 12 | 8 | 16 | 24 | 56 | 12 | 28 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 8 | 12 | 26 | 38 | 58 | 78 | 68 | 52 | 45 | 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 12 | 16 | 32 | 60 | 72 | 120 | 192 | 104 | 188 | 68 | 152 | 24 | 52 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 8 | 16 | 36 | 84 | 136 | 166 | 296 | 314 | 322 | 350 | 350 | 274 | 232 | 180 | 112 | 80 | 34 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 20 | 24 | 40 | 80 | 136 | 292 | 332 | 600 | 640 | 684 | 804 | 684 | 1024 | 488 | 852 | 264 | 536 | 132 | 296 | 40 | 84 | 0 | 0 |
| 12 | 8 | 20 | 52 | 104 | 186 | 330 | 524 | 704 | 1022 | 1348 | 1544 | 1790 | 1933 | 1896 | 1908 | 1806 | 1636 | 1410 | 1138 | 904 | 654 | 480 | 340 |
| 13 | 24 | 24 | 52 | 132 | 228 | 376 | 608 | 1040 | 1248 | 2232 | 2376 | 3524 | 3488 | 4484 | 4452 | 4664 | 5268 | 3992 | 5196 | 2880 | 4368 | 1876 | 3308 |
| 14 | 12 | 40 | 68 | 132 | 252 | 464 | 778 | 1290 | 1916 | 2624 | 3736 | 4992 | 6210 | 7670 | 9000 | 10234 | 11214 | 11886 | 12186 | 11818 | 11436 | 10788 | 9704 |
| 15 | 16 | 40 | 84 | 152 | 284 | 524 | 904 | 1444 | 2484 | 3580 | 4948 | 7380 | 8936 | 12700 | 13674 | 19420 | 19184 | 26108 | 24908 | 30180 | 29224 | 30628 | 31308 |
| 16 | 16 | 24 | 60 | 176 | 340 | 596 | 1072 | 1774 | 2808 | 4432 | 6600 | 9286 | 13138 | 17410 | 22406 | 28488 | 34735 | 41928 | 49332 | 56258 | 62858 | 68250 | 75080 |
| 17 | 32 | 48 | 80 | 184 | 356 | 708 | 1180 | 1992 | 3140 | 4904 | 8224 | 11700 | 17832 | 23620 | 32384 | 41108 | 52236 | 67324 | 77952 | 101244 | 106908 | 138092 | 136448 |
| 18 | 12 | 36 | 108 | 204 | 392 | 712 | 1272 | 2208 | 3816 | 6118 | 9382 | 14524 | 21072 | 29640 | 41840 | 56552 | 74910 | 97174 | 123214 | 152678 | 184688 | 219860 | 256646 |
| 19 | 36 | 32 | 92 | 240 | 448 | 840 | 1424 | 2612 | 4108 | 7080 | 10564 | 16408 | 25160 | 36272 | 53612 | 72332 | 102576 | 128732 | 177452 | 214540 | 281348 | 336756 | 415164 |
| 20 | 16 | 40 | 88 | 216 | 472 | 884 | 1632 | 2824 | 4704 | 7694 | 12434 | 19224 | 29104 | 43590 | 63098 | 89990 | 125784 | 171134 | 229774 | 301350 | 388639 | 492738 | 614798 |

**Figure 15.** In row $L$, column $K$ there are the number of (undirected) free homotopy classes of closed curves in the **punctured torus** of word length $L$ and self-intersection $K$

The following fact is a direct consequence of Theorem 5.1, [14, Corollary 1.8].

**Theorem 5.2.** *The pair of pants is the only surface with negative Euler characteristic that has only finitely many free homotopy classes of curves with a given self-intersection.*

**Theorem 5.3.** *The maximal self-intersection number of a free homotopy class (possibly a power) in the pair of pants of word length L is bounded above by* $\frac{L^2}{4} + \frac{L}{2} - 1$. *If L is even this bound is sharp.*

*The maximal self-intersection of a non-power free homotopy class in the punctured torus of word-length L is* $\lfloor \frac{L^2-2}{4} \rfloor$. *This bound is sharp.*

## 5.2 *Conjectures*

Table 14 lead us to the following:

**Conjecture 1.** *In the pair of pants there is an increasing sequence, starting with*

$$2, 6, 12, 20, 32, 52, 92, 156, 244, 360, 524 \ldots$$

*such that for each L in* $\{K+3, K+4, \ldots, 2K+1\}$, $\mathrm{P}(K, L)$ *is the* $(2K+2-L)$*-th term of the sequence.*

We observed the following patterns in the entries $\mathrm{P}(K, L)$ of these tables, and analogous tables we computed for other surfaces with boundary.

**Conjecture 2.** *For all surfaces with boundary*

*(1) If two cells in the same row have a positive entry, all the cells in between in that same row also have positive entries.*

*(2) The sequence* $\{a_K\}$ *defined by*

$$a_K = \min\{\mathrm{WL}(w) : \mathrm{SI}(w) = K\}$$

*is increasing.*

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| 12 | 0.11 | 0.28 | 0.72 | 1.44 | 2.58 | 4.58 | 7.28 | 9.78 |
| 13 | 0.28 | 0.28 | 0.62 | 1.56 | 2.70 | 4.45 | 7.20 | 12.31 |
| 14 | 0.12 | 0.41 | 0.69 | 1.35 | 2.57 | 4.73 | 7.94 | 13.16 |
| 15 | 0.14 | 0.36 | 0.75 | 1.35 | 2.52 | 4.66 | 8.04 | 12.84 |
| 16 | 0.13 | 0.19 | 0.47 | 1.38 | 2.66 | 4.66 | 8.38 | 13.86 |
| 17 | 0.22 | 0.33 | 0.55 | 1.27 | 2.46 | 4.90 | 8.17 | 13.79 |
| 18 | 0.07 | 0.22 | 0.67 | 1.26 | 2.42 | 4.40 | 7.85 | 13.63 |
| 19 | 0.20 | 0.18 | 0.51 | 1.33 | 2.48 | 4.65 | 7.89 | 14.47 |
| 20 | 0.08 | 0.20 | 0.44 | 1.08 | 2.36 | 4.42 | 8.16 | 14.12 |
| 21 | 0.11 | 0.29 | 0.58 | 1.20 | 2.27 | 4.55 | 7.95 | 14.11 |

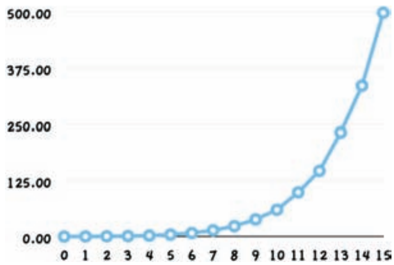


**Figure 16.** On the left: in row $L$, column $K$ there are the number of free homotopy classes of closed curves in punctured torus of word length equal to $L$ and self-intersection $K$, all divided by $L^2$. On the right: The function $K \mapsto \mathrm{P}(K, 21)$ for $K \in \{0, 1, \ldots, 15\}$

*(3) Denote by g and b the genus and the number of boundary components of a surface with negative Euler characteristic. For each $K$, the sequence $\mathrm{P}(K, L)/L^{6g+2b-6}$ converges to a positive number $c_K$ as $L$ goes to infinity. The sequence $\{c_K\}$ is exponentially increasing. (see figure 16).*

It is not hard to see that (1) implies (2) in Conjecture 2. Compare Conjecture 2(3) with [29].

We have also computed word-length/self-intersection tables for other surfaces with boundary. These experiments did not suggest precise polynomials as in Theorem 5.3, but they lead to the following conjecture (see [14, Conjecture 1.10]).

**Conjecture 3.** *Consider a surface $\Sigma$ with boundary and negative Euler characteristic $\chi$. Denote by $\mathrm{SI}_{\max}(L)$ the maximum self-intersection number for all free homotopy classes of closed curves on $\Sigma$ of word length at most $L$. Then*

$$\lim_{L\to\infty} \frac{\mathrm{SI}_{\max}(L)}{L^2} = \frac{\chi}{2\chi - 1}.$$

*Remark 5.4.* Consider a representative of a class of word length $L$ that can be decomposed as a union of $L$ straight segments in the polygon obtained by removing from the surface the arc chosen in Subsection 2.1. Since lines intersect in at most one point, the maximal number of self-intersection points is less than or equal to $\binom{L}{2} = \frac{L(L-1)}{2}$. On the other hand, when the Euler characteristic goes to infinity, $\lim_{L\to\infty} \frac{\mathrm{SI}_{\max}(L)}{L^2} = \frac{\chi}{2\chi-1}$ in Conjecture 3 goes to $\frac{1}{2}$. This is a consistency check.

## 6. Growth of the number of geodesics with given self-intersection

We know, by Theorem 3.7 that the growth of the number of geodesics up to length $L$ is exponential on $L$. In this section we partition these into those with given self-intersection number.

### 6.1 *Growth of the number of simple geodesics*

For each non-negative integer $K$, denote by $\mathcal{C}_\Sigma(L, K)$ the cardinal of the set of geodesics in $\Sigma$ of self-intersection number $K$ and geometric length at most $L$.

For a long time, many researchers tried to determine the growth of $\mathcal{C}_\Sigma(L,0)$. In 1985, Birman and Series [5] proved that the union of all closed geodesics covers a very "thin" set.

**Theorem 6.1.** *For each non-negative integer k, the set $S_k$ of points on a hyperbolic surface which lie on a complete geodesic of self-intersection number at most k is nowhere dense and has Hausdorff dimension one.*

For simple curves Theorem 6.1 was observed first by Bill Thurston in the mid seventies.

Birman and Series stated that the number of simple geodesics of length at most $L$ is bounded by a polynomial of degree $6g+2b-6$, where $g$ and $b$ are the genus and number of boundary components of the surface respectively.

In 2001, Rivin [33] proved the following:

**Theorem 6.2.** *Let $\Sigma$ be a hyperbolic surface of genus g, b boundary components and c cusps. Then there exist constants $c(\Sigma)$ and $d(\Sigma)$ such that*

$$c(\Sigma)\cdot L^{6g+2b+2c-6} \le \mathcal{C}_\Sigma(L,0) \le d(\Sigma)\cdot L^{6g+2b+2c-6}.$$

In her Ph.D. Thesis [28], (see also [29]) Mirzakhani proved the following

**Theorem 6.3.** *Let $\mathcal{M}_{g,n}$ be the moduli space of complete hyperbolic Riemann surfaces of genus g with n cusps. Then*

$$\mathcal{C}_\Sigma(L,0) \sim c(\Sigma)L^{6g+2n-6}$$

*where $c\colon \mathcal{M}_{g,n} \longrightarrow \mathbb{R}$ is a continuous proper function. (Recall that $\sim$ means asymptotic).*

Mirzakhani also computes the leading coefficient.

Theorem 6.3 is a part of Mirzakhani's Field's medal work. Another concerns flows in Teichmuller spaces of all hyperbolic surfaces with a fixed topology, see for example [1].

*Example 3.* We illustrate Theorem 6.3 as follows. In the rectangular torus (14, 16) described in Appendix A.2 we considered the set of all geodesics of geometric length up to 280 as in Example 1. From this set, we selected, using algorithm described in Subsection 4.1, all simple closed geodesics and computed the geometric length of each of them. In this metric, there are 225 simple geodesics up to length 280. We order the lengths of these 225 geodesics from smaller to larger: $\{l_1, l_2, \ldots, l_{225}\}$.

In figure 17 we plotted the points of the form $(i, l_i)$ for each $i$ in $\{1, 2, \ldots, 225\}$.

Define a function $F(L) = \mathcal{C}_\Sigma(L,0)$, that is $F(L)$ is the number of simple geodesics in our hyperbolic torus of length at most $L$. By Mirzakhani's Theorem 6.3, this function is asymptotic to $c(\Sigma)L^2$. If we "pretend" that $F(L) = c(\Sigma)L^2$, then $F^{-1}(i) = l_i$, since $i$ is the number of simple geodesics of length at most $l_i$. Thus the set of points of the form $(i, l_i)$ $\{l_1, l_2, \ldots, l_{225}\}$ can be approximated by a function $g(x) = d\cdot\sqrt{x}$ for some constant $d$. By least squares, we approximated this constant, obtaining 18.53. The function $g(x) = 18.53\sqrt{x}$ together with the points is plotted in figure 17.

Observe that Mirzakhani's result is asymptotic, thus it is not a "given" that one can get a nice approximation as the one displayed in figure 17.

*Remark 6.4.* The pair of pants has finitely many geodesics of each self-intersection. Therefore, an analysis similar to that of Example 3 will not give any interesting output.

### 6.2 *Growth of the number of geodesics with self-intersection larger than zero*

Using Mirzakhani's approach, Rivin [34] proved:

**Theorem 6.5.** *There exists a constant $c(\Sigma)$, depending on the hyperbolic structure on $\Sigma$ such that:*

$$\mathcal{C}_\Sigma(L, 1) \sim c(\Sigma)L^{6g+2n-6}$$

*Example 4.* We repeated the computations we made in Example 3 but now considering geodesics of self-intersection one. There are 329 geodesics of self-intersection one and length smaller than 280, see figure 17.

We also repeated the computations we made in Example 3 considering geodesics of self-intersection number 10, 20, 30, 40, 50, 60 and 70. The results are displayed in figure 18. Observe that in that figure, the scale of the $x$-axis varies considerably, since the number of geodesics of fixed self-intersection number up to a certain length, also varies considerably (compare to Subsection 8.2). Moreover, by Theorem 7.5, a geodesic with "large" self-intersection number, say 70, has to be "long enough". In our example, there are about 600 closed geodesics of self-intersection number 70 and geometric length at most 280. The shortest one has length approximately 279.227.

The previous results, as well as computer experiments, suggest that the growth of the number of geodesics of a given self-intersection is the same as the growth of simple geodesics (with a different leading coefficient).

Fix a hyperbolic surface $\Sigma$ and denote by $G(K, L)$ the number of closed geodesics of self-intersection number $K$ and geometric length at most $L$.

**Conjecture 4.** *For each hyperbolic surface of genus $g$, $b$ geodesic boundary components and $p$ punctures and for each $K$, the sequence $G(K, L)/L^{6g+2b+2p-6}$ converges to a positive number $d_K$ as $L$ goes to infinity. The sequence $\{d_K\}$ is exponentially increasing.*

For information about this conjecture see the preprints [36] and [35] by Sapir, a former student of Mirzakhani.

*Remark 6.6.* Assuming the graphs of figure 18 have the form $y = c\sqrt{x} + d$, then $c$ can be estimated as $2\sqrt{x_0}$ where $x_0$ is the $x$-coordinate where the slope changes from $+\infty$ to 0, (say equals 1). After $x_0$ the graph looks like a horizontal line in this re-scaling.

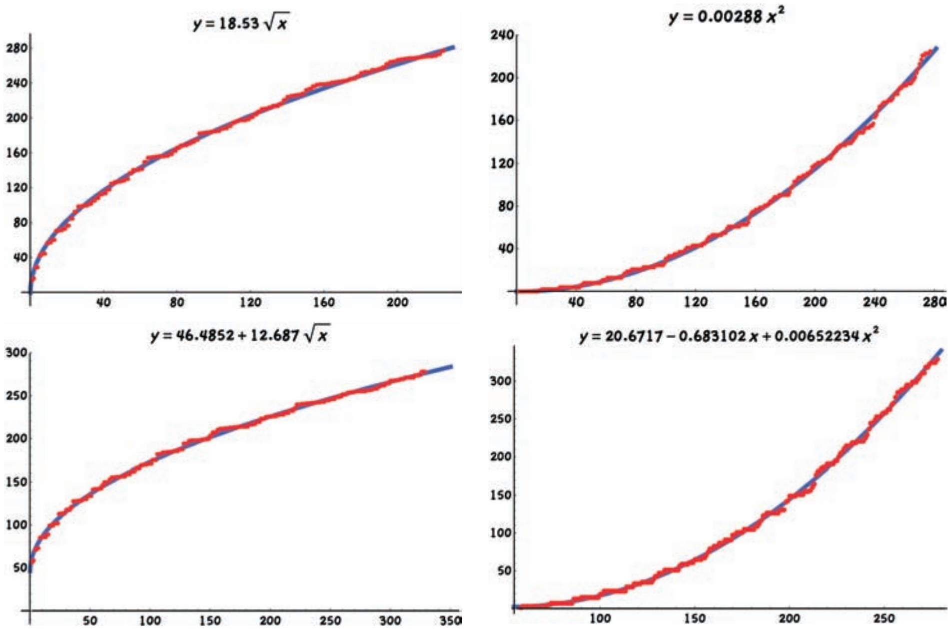


**Figure 17.** Top left: The points of the form $(i, l_i)$ from Example 3 and the graph of the function $g(x) = 18.53\sqrt{x}$. Top right: The points of the form $(n, \mathcal{C}_\Sigma(n, 0)$ and the function $f(x) = 0.0028x^2$. Bottom left: The points of the form $(i, l_i)$ from Example 4 and the graph of the function $g(x) = 46.4852 + 12.687\sqrt{x}$. Bottom right: The points of the form $(n, \mathcal{C}_\Sigma(n, 1)$ and the function $f(x) = 20.6717 - 0.683102x + 0.00652234x^2$.

## 7. Relations between geometric length and self-intersection number

Recall that by Theorems 3.7 and 3.9, the growth of the number of closed geodesics in a hyperbolic surface up to geometric length $L$ is exponential in $L$. On the other hand, the growth of the number of simple geodesics in hyperbolic surfaces up to length $L$ is polynomial in $L$ by Theorem 6.3. In the pair of pants there are only three unoriented simple geodesics (the corresponding polynomial has degree zero). For all the rest of the hyperbolic surfaces the polynomial has positive degree. Thus, in these cases, there is no upper bound on the length of simple geodesics. Conjecture 4 states that an analogous result holds for geodesics of a fixed self-intersection number. In this section, we will explore lower bounds on the geometric length of geodesics of a given self-intersection. Roughly speaking, a geodesic has to be "long" to have a large number of self-intersections. Moreover Basmajian [3] proved that the number of self-intersections of a closed geodesic is bounded by a constant times the square of the hyperbolic length. Compare to Theorem 5.3 and Conjecture 3.

**Theorem 7.1.** *If $\Sigma$ is a hyperbolic surface with (possibly empty) geodesic boundary then there exists a constant $d(\Sigma)$ depending on the hyperbolic metric on $\Sigma$ such that* $\mathrm{SI}(w) \leq d(\Sigma)\,\mathrm{WL}(w)^2$ *where $d(\Sigma)$ is a continuous function of $\Sigma$.*

*Example 5.* To illustrate Theorem 7.1, we computed the tables displayed in figure 19, where for some geometric lengths $L$, we computed the maximum of $\mathrm{SI}(w)/\,\mathrm{GL}(w)^2$

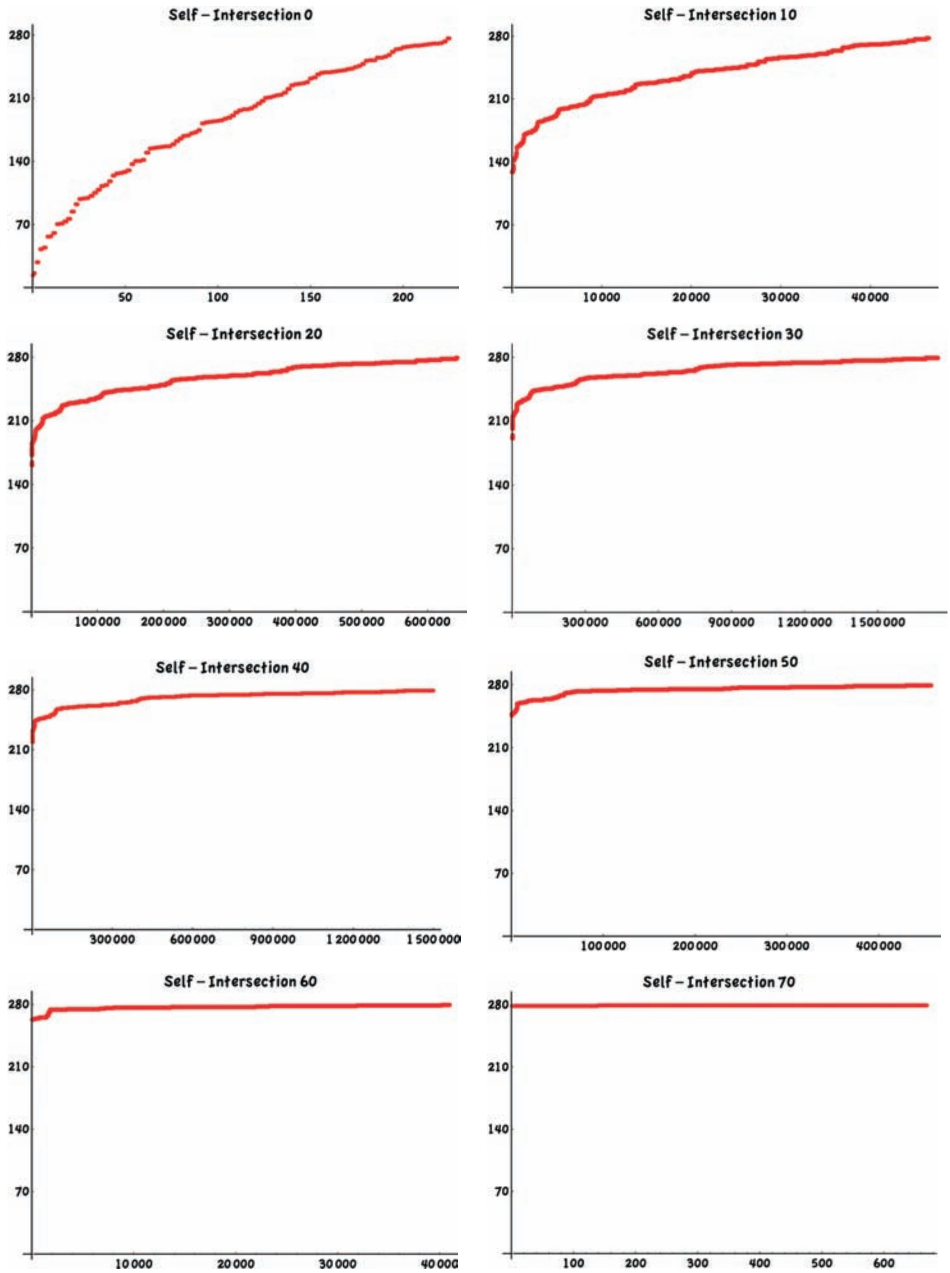


**Figure 18.** For each $s \in \{0, 10, \ldots 70\}$, we order (from smallest to largest) the lengths of the $n(s)$ geodesics of self-intersection $s$ and geometric length at most 280, $\{l_1, l_2, \ldots, l_{n(s)}\}$ and plot points of the form $(i, l_i)$

for all $w$ such that $\mathrm{GL}(w) < L$. If $\Sigma$ is the hyperbolic pair of pants $(2.5, 2.6, 5)$ as in Definition A.4 then $d(\Sigma)$ is approximately 0.025 and in the case of the rectangular torus (14, 16) (Definition A.7), the constant seems to be approximately 0.0009.

**Pants 2.5, 2.6, 5**

| up to GL | Max(SI(w)/GL(w)²) |
|---|---|
| 22.5 | 0.024709423 |
| 25 | 0.024952905 |
| 27.5 | 0.024952905 |
| 30 | 0.024952905 |
| 32.5 | 0.025073507 |
| 35 | 0.025073507 |
| 37.5 | 0.025073507 |
| 40 | 0.025073507 |
| 42.5 | 0.025073507 |
| 45 | 0.025073507 |

**Rectangular torus 14-16**

| up to GL | Max(SI(w)/GL(w)²) |
|---|---|
| 126.0 | 0.000654914 |
| 140.0 | 0.000696838 |
| 154.0 | 0.000738147 |
| 168.0 | 0.000769400 |
| 182.0 | 0.000795936 |
| 196.0 | 0.000820391 |
| 210.0 | 0.000838317 |
| 224.0 | 0.000858166 |
| 238.0 | 0.000870696 |
| 252.0 | 0.000887266 |

**Figure 19.** Estimation of the constant $d(\Sigma)$ of Theorem 7.1

## 7.1 *Lower bounds for the length of the shortest geodesic of a given self-intersection*

A geodesic (in any hyperbolic surface, possibly with puntures and geodesic boundadry) shorter than $4\log(1+\sqrt{2})$ is simple, [22,31,42,43]. In [2] and [3] Basmajian proved the following generalization:

**Theorem 7.2.** *For all compact hyperbolic surfaces with geodesic boundary and finitely many cusps, there exists a sequence of constants $\{M_k\}$ going to infinity with $k$ such that if a geodesic in $w$ has geometric length smaller than $M_k$ then* $\mathrm{SI}(w) < k$.

Basmajian extended these results in [3], showing that these constants are sharp, in the following sense:

**Theorem 7.3.** *For each $k \in \mathbb{N}$, there exists a hyperbolic surface and a geodesic on that surface with self-intersection number $k$ and geometric length $M_k$. Moreover,*

$$\frac{\log(2k)}{4} \leq M_k \leq 2\cosh^{-1}(2k+1).$$

As a corollary Basmajian stated,

**Theorem 7.4.** *If $w$ is a closed geodesic on a hyperbolic surface (possibly with punctures and geodesic boundary) satisfying* $\mathrm{GL}(w) \leq \frac{\log(2k)}{4}$ *then the self-intersection number of $w$ is at most $k-1$.*

Basmajian asked if the following were true.

**Conjecture 5.** *The sequence $\{M_k\}$ is increasing.*

Observe that the constants $M_k$ are universal in the sense that they hold for any hyperbolic surface (possibly with punctures and geodesic boundary). One can ask similar questions for a fixed hyperbolic surface: Consider the sequence $\{s_k\}_{k\geq 0}$, where $s_k$ is the geometric length of the shortest geodesic of self-intersection number $k$. It is not hard to prove that $s_0 < s_k$ for all $k > 0$. Buser [7, Theorem 4.2.4] proved that in a closed hyperbolic surface, $s_1 < s_k$ for all $k > 2$.

Basmajian proved.

**Theorem 7.5.** *Let $\Sigma$ be a hyperbolic surface.*

*(1) Let $\Sigma$ be a compact surface with (possibly empty) geodesic boundary. Denote by $L(\Sigma)$ the length of the shortest geodesic with one self-intersection on $\Sigma$. Then there exists a constant $c(\Sigma)$ so that,*

$$c(\Sigma)\sqrt{k} \leq s_k \leq 3L(\Sigma)(\sqrt{k}+1),$$

*Moreover, the constant $d(\Sigma)$ of Theorem 7.1 satisfies $d(\Sigma) = 1/c(\Sigma)^2$.*

*(2) If $\Sigma$ has at least one cusp and is not the punctured disc, then for $k = 2, 3, \ldots$*

$$\frac{1}{2}\log\left(\frac{k}{2}\right) \leq s_k(\Sigma) \leq 2\sinh^{-1}(k) + d(\Sigma) + 1$$

*where $d(\Sigma)$ is the shortest orthogonal distance from the length one boundary of a cusp in $\Sigma$ to itself.*

**Conjecture 6.** *The sequence $\{s_k\}$ is increasing. Moreover, for each hyperbolic surface $\Sigma$ of genus $g$ and $b$ (geodesic) boundary components there exists a constant $u(\Sigma)$ (depending on the metric) such that the sequence $\{s_k - u(\Sigma)\sqrt{k}\}$ is bounded.*

*Moreover, in the case of the pair of pants the sequence $\{s_k - u(\Sigma)\sqrt{k}\}$ approaches zero when $k$ goes to infinity, and in the other hyperbolic surfaces, the sequence approaches a positive constant.*

Observe that Conjecture 2(1) is the combinatorial version of Conjecture 6.

Our computer evidence supports Conjecture 6. We tested many different metrics in the pair of pants and many metrics in the rectangular torus with one boundary component described in Subsection A.2.

*Example 6.* The length of the shortest geodesic with one-self-intersection in the rectangular torus $(14, 16)$ (Definition A.7), $L(\Sigma)$, is approximately 57.2274. In Example 5 we estimated that constant $d(\Sigma)$ is approximately 0.0009. By Theorem 7.5, $c(\Sigma)$ is approximately 33.33.

$$33.33\sqrt{k} \leq s_k \leq 171.6822(\sqrt{k}+1)$$

In the pair of pants $(2.5, 2.6, 5)$, $d(\Sigma)$ is approximately 0.025. Thus, $c(\Sigma)$ is about 6.32. The length $L(\Sigma)$ of the shortest geodesic with one self-intersection is 6.6.

$$6.32\sqrt{k} \leq s_k \leq 19.8(\sqrt{k}+1)$$

## 8. Statistics relating geometric length, word length and self-intersection

### 8.1 *Self-intersection sampling by word-length*

Figure 21 shows the histogram of all non-power, free homotopy classes in the punctured torus of word length $4, 6, 8, \ldots, 20$, organized by self-intersection. The histograms are depicted up to scale. They suggest that the distribution of self-intersection sampling by word-length (appropriately normalized) converges to a Gaussian when the word length goes to infinity. We proved this result jointly with Steve Lalley [12]:

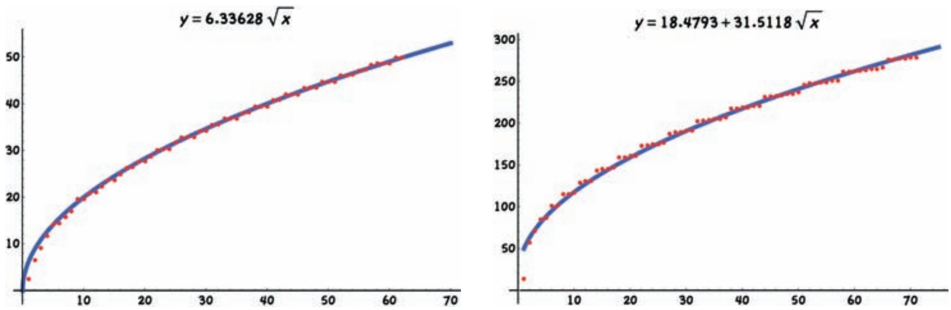


**Figure 20.** Two graphs of the sequence $\{s_k\}$ of Conjecture 6 (as a function of $k$) for the pair of pants (2.5, 2.6, 5) (left) and on the right, a rectangular torus (14, 16)

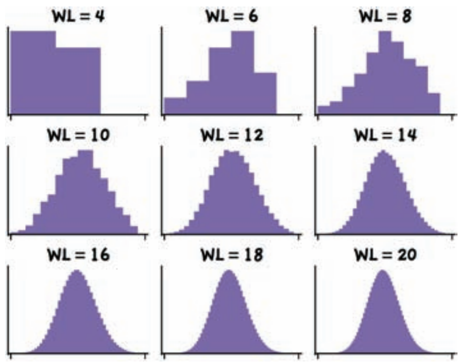


**Figure 21.** Histograms of free homotopy classes of closed curves in the punctured torus, organized by word length

**Theorem A.** *On a surface with non-empty boundary and negative Euler characteristic $\chi$, the proportion of words $w$ of word length $L$ such that*

$$a < \frac{\mathrm{SI}(w) - \kappa \cdot L^2}{\sigma L^{3/2}} < b$$

*converges as $L$ goes to infinity to* $\frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-\frac{x^2}{2}}\, dx$ *where*

$$\kappa = \frac{\chi}{3(2\chi - 1)} \quad \textit{and} \quad \sigma^2 = \frac{2\chi(2\chi^2 - 2\chi + 1)}{45(2\chi - 1)^2(\chi - 1)}.$$

*In other words, when $L$ is very large, the distribution of self-intersection of all free homotopy classes of word length $L$ is close to a Gaussian with mean $\kappa \cdot L^2$ and standard deviation $\sigma \cdot L^{3/2}$.*

*Remark 8.1.* Observe that with word length as small as 20 we obtain a strongly Gaussian-like histogram. Of course, the population in this case is extremely large).

*Remark 8.2.* The expected value of intersections of $n$ random chords in a circle is $\frac{n(n-1)}{6}$ and the variance is $\frac{n(n-1)(n+3)}{45}$, [38, Chapter 6] (Here, a random chord is determined by two points independently and randomly placed on the circumference, with uniform distribution). Compare to the above: when the Euler characteristic $\chi$ is very large, the mean of self-intersections of all classes of word length $L$ is close to $\frac{L^2}{6}$ when $L$ is large and the variance is close to $\frac{L^3}{45}$.

*Remark 8.3.* In the case of the pair of pants and torus with one boundary component, the mean of self-intersection of all classes of word length $L$ is approximately $\frac{L^2}{9}$ for large $L$, while the maximal self-intersection is about $\frac{L^2}{4}$.

## 8.2 *Self intersection sampling by geometric length*

Lalley [24] and [25] proved.

**Theorem 8.4.**

*(1) If $\Sigma$ is a closed surface of genus $g \geq 2$ and constant negative curvature c, then for L large, "most" closed geodesics of geometric length at most L have self-intersection close to $\frac{c}{2\pi(g-1)}L^2$.*
*(2) Let $\Sigma$ be a closed hyperbolic surface. Denote by $\omega_L$ a random closed geodesic chosen among all geodesics of geometric length at most L, then for some probability distribution $\Phi$, and some constant $\kappa = \kappa(\Sigma)$, $\frac{\mathrm{SI}(\omega_L)-\kappa L^2}{L}$ converges in distribution to $\Phi$.*
*(3) If $\Sigma$ is a closed surface with variable negative curvature, then for some constants $\kappa' = \kappa'(\Sigma)$ and $\sigma$, $\frac{\mathrm{SI}(\omega_L)-\kappa L^2}{\sigma L^{3/2}}$ converges in distribution to the standard unit Gaussian distribution.*

Our experiments indicate that for surfaces with geodesic boundary, a result analogous to Theorem 1.1 (sampling by word length) and Lalley's theorem 8.4 for variable curvature (sampling by geometric length) holds, figure 22 depicts the output of some of our experiments.

**Conjecture 7.** *For a hyperbolic surface with geodesic boundary the distribution of self-intersection sampling by geometric length (appropriately normalized) converges (in distribution) to a Gaussian, when the geometric length goes to infinity.*

## 8.3 *Geometric length sampling by word length*

In [13] we conjectured that the geometric length, sampling by word length approaches a Gaussian when the word length goes to infinity, see figure 23. Conversely, computations also suggest that the distribution of word length sampling by geometric length tends to a Gaussian when the geometric length goes to infinity.

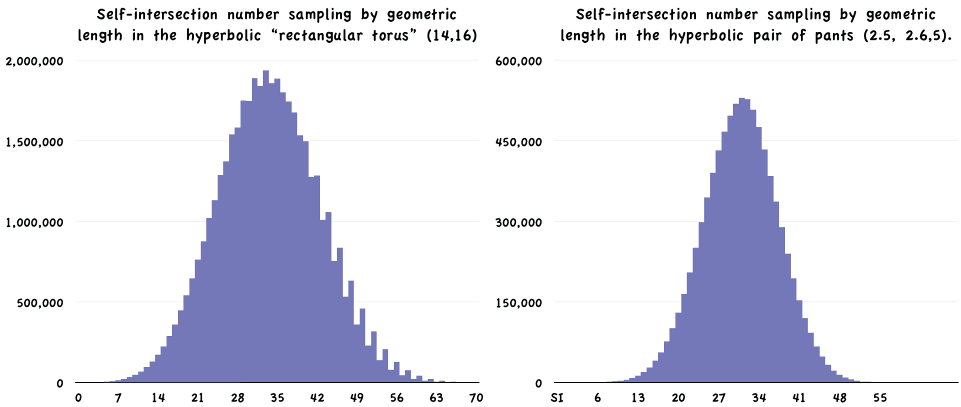


**Figure 22.** Histogram of the self-intersection number of all geodesics up to length 38 in the hyperbolic pair of pants of (2.5, 2.6, 5) on the left. On the right histogram of the intersection of all geodesics up to geometric length 280 in the rectangular torus h (14, 16)

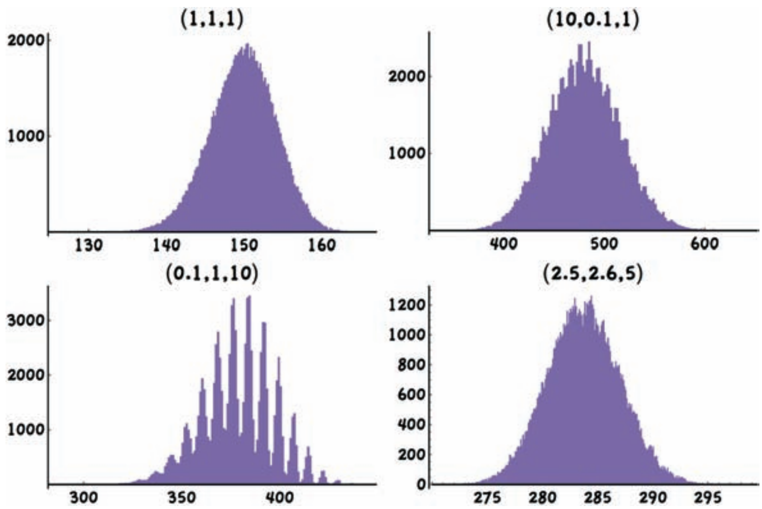


**Figure 23.** Histograms of the geometric length of a sample of 100,000 words of word length 100 in different hyperbolic pair of pants. The parameters $(A, B, C)$ of the length of the boundary components are indicated in each graph

## A. How we sample by geometric length

### A.1 *The pair of pants*

By [7, Theorem 2.4.2], given three positive real numbers $\alpha, \beta$ and $\gamma$, there exists a unique triple of numbers $u, v$ and $w$ and a unique, convex, right angled hyperbolic hexagon with side lengths $\alpha, u, \beta, v, \gamma, w$ taken in cyclic order.

Given three positive real numbers $la, lb$ and $lc$, set $\alpha = la/2$, $\beta = lb/2$ and $\gamma = lc/2$, consider two congruent hexagons of alternating side lengths $\alpha, \beta$ and $\gamma$,

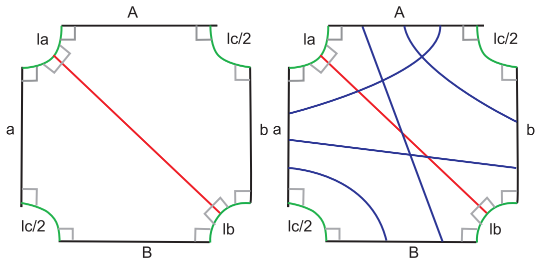


**Figure 24.** The octogon obtained "cutting up" a hyperbolic pair of pants on the left and a representative of the curve $aaBab$ on the right

glue the obvious sides to obtain a hyperbolic pair of pants of geodesic with boundary components of lengths $la$, $lb$ and $lc$, figure 24.

**Proposition A.1.** *Let* $C = \min\{la, lb, lc/2\}$. *For each geodesic* $w$ *in the pair of pants with boundary components of length* $la, lb, lc$,

$$\mathrm{GL}(w) \geq C \cdot \mathrm{WL}(w),$$

*where the word length is computed in the* $a, b$*-alphabet.*

*Proof.* A closed geodesic $w$ can be decomposed as the union of geodesic segments with endpoints in the edges labeled $a$, $A$, $b$ and $B$.

We claim that each of these segments is longer than $C$. Thus, the result follows.

We will prove this claim in the case of the geodesic $aaBab$ depicted in figure 24.

Consider the segment from the $B$-edge, to the $a$-edge. This segment is the top side of a quadrilateral with two right angles at the base. Thus it is longer than the base, which has length $lc/2$.

Now, let's consider the segment from the $A$-edge to the $B$-edge. This segment is the union of the segment $x_1$ from the $A$-edge to the segment from $la$ to $lb$ (in red in figure 24) and $x_2$ from the segment from $la$ to $lb$ to the $B$-edge.

Analogously as before, $x_1$ is longer than $la/2$ and $x_2$ is longer than $lb/2$. The proof of the claim for the other segments, and for the general case can be completed with the same ideas. □

**Corollary A.2.** *In the pair of pants with boundary components of length* $la, lb, lc$,

$$\{w \in \pi_0 : \mathrm{GL}(w) \leq L\} \subset \{w \in \pi_0 : \mathrm{WL}(w) \leq C \cdot L\},$$

*where* $C = \min\{la, lb, lc/2\}$ *and the word length is computed in the* $a, b$*-alphabet.*

*Remark A.3.* The constant $\min\{la, lb, lc/2\}$ is sharp since $\mathrm{GL}(a) = la$, $\mathrm{GL}(b) = lb$ and $\mathrm{GL}(ab) = lc$.

**Definition A.4.** *A* pair of pants $(la, lb, lc)$ *is a pair of pants with a hyperbolic metric and geodesic boundary components of length* $la$, $lb$ *and* $lc$.

We will now describe the metric of the particular pair of pants we discussed in this work. For details see [13].

Set

$$\alpha = \begin{pmatrix} \cosh\left(\frac{la}{2}\right) & \cosh\left(\frac{la}{2}\right) - 1 \\ \cosh\left(\frac{la}{2}\right) + 1 & 1 \end{pmatrix} \text{ and } \beta = \begin{pmatrix} \cosh\left(\frac{lb}{2}\right) & \Delta \\ \sinh^2\left(\frac{lb}{2}\right)/\Delta & \cosh\left(\frac{lb}{2}\right) \end{pmatrix}$$

where $\Delta = -\dfrac{\cosh\left(\frac{la}{2}\right)\cdot\cosh\left(\frac{lb}{2}\right)+\cosh\left(\frac{lc}{2}\right)+\sqrt{\frac{1}{2}\left(4\cosh\left(\frac{la}{2}\right)\cosh\left(\frac{lb}{2}\right)\cosh\left(\frac{lc}{2}\right) + \cosh(la)+\cosh(lb)+\cosh(lc)+1\right)}}{\cosh\left(\frac{la}{2}\right)+1}$.

The matrices $\alpha$ and $\beta$ generate a discrete, cocompact subgroup $G$ of $PSL(2,\mathbb{R})$. Recall that a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $PSL(2,\mathbb{R})$ acts on the upper half plane $\mathbb{H}$ by $z \mapsto \frac{a\cdot z+b}{c\cdot z+d}$. The quotient of $\mathbb{H}$ by $G$ is our pair of pants. The right angled hyperbolic octogon of figure 24 is a fundamental domain of the action of $G$ on $\mathbb{H}$. The quotient map $\mathbb{H} \longrightarrow \mathbb{H}/G$ is a covering map and the action of $G$ on $\mathbb{H}$ is given by the deck transformations.

Given a cyclic reduced word $w$ in the generators $a, b$ of the fundamental group of the pair of pants (and their inverses), if one replaces each occurrence of $a$ by the matrix $\alpha$, $a^{-1}$ by the matrix $\alpha^{-1}$ and similarly with $b$ and $\beta$, one obtains a matrix $g$. The length of the closed geodesic in $w$ satisfies:

$$\cosh\left(\frac{\mathrm{GL}(w)}{2}\right) = \frac{|\mathrm{tr(g)}|}{2}.$$

Thus, one can program the computer to find the length of all closed geodesics using the matrices $\alpha$ and $\beta$. In the example of the pair of pants (2.5, 2.6, 5) the matrices are (approximately)

$$\alpha = \begin{pmatrix} 1.888 & 0.888 \\ 2.888 & 1.888 \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} 1.970 & -6.690 \\ -0.431 & 1.970 \end{pmatrix}$$

In order to study all geodesics up to geometric length 50 in the pair of pants (2.5, 2.6, 5), we have our program compute the geometric length of all cyclic reduced words up to word length 20, and choose from this set those with geometric length at most 50. By Corollary A.2, this is enough.

Observe that Theorem 3.14 would not have sufficed: Theorem 3.14 implies that there exists a constant $C$ such that the inclusion of Corollary A.2 holds, but does not say explicitly *what* the constant is.

### A.2 *The torus with one boundary component*

In this Subsection we describe a metric on the torus with one geodesic boundary component due to Bernard Maskit, which satisfies a statement similar to Proposition A.1. This will allow us to study all geodesics up to a given length.

Fix two positive numbers $la$ and $lb$ such that $\sinh la \cdot \sinh lb > 1$. Consider two perpendicular geodesic segments with length $la/2$ and $lb/2$. Drop perpendiculars to each of these segments from the endpoint the segments do not have in common. By [7, Lemma 2.3.5], these two lines are disjoint, and moreover, have a common perpendicular. Hence, a right-angled pentagon as in figure 25 (right) is obtained.

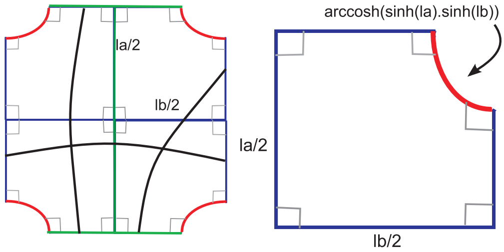


**Figure 25.** Left: Arcs of geodesics in a "rectangular torus". Right: One of the four right angled pentagons that form a rectangular torus

By gluing the appropriate pairs of congruent edges of four of these pentagons we obtain a hyperbolic torus with geodesic boundary.

Using the same type of ideas of the proof of Proposition 24, and the fact [7, Lemma 2.3.5] that the length of the side of the right angle pentagon is $\operatorname{arccosh}(\sinh la \sinh lb)$ one can prove:

**Proposition A.5.** *For each geodesic $w$ in Maskit's torus with one boundary component and length of generators $la, lb$,*

$$\mathrm{GL}(w) \geq C \cdot \mathrm{WL}(w),$$

*where $C = \min\{la, lb, lc\}$ and $\cosh lc = \sinh la \sinh lb$. The word length is computed in the $a, b$-alphabet. Therefore,*

$$\{w \in \pi_0 : \mathrm{GL}(w) \leq L\} \subset \{w \in \pi_0 : \mathrm{WL}(w) \leq C \cdot L\},$$

*Proof.* A geodesic of word length $L$ can be decomposed in $N$ segments that either run "parallel" to one of the generators or parallel to one fourth of the boundary component. Thus, each is at least $\min\{la, lb, bd/4\}$, where $la$ and $lb$ denote the length of the two generators and $bd$ the length of the boundary component. □

*Remark A.6.* The constant $\min\{la, lb, bd/4\}$ is sharp since $\mathrm{GL}(a) = la$, $\mathrm{GL}(b) = lb$ and $\mathrm{GL}(abAB) = lc$.

**Definition A.7.** *A* rectangular torus $(la, lb)$ *is a torus with a hyperbolic metric and generators of length la and lb, constructed gluing the appropriate pairs of edges of four copies of a right angled pentagon as in figure 25.*

Consider the matrices

$$\alpha = \begin{pmatrix} e^l a & 0 \\ 0 & e^{-la} \end{pmatrix} \text{ and } \beta = \begin{pmatrix} \cosh(lb) & -\sinh(lb) \\ -\sinh(lb) & \cosh(lb) \end{pmatrix}.$$

We compute lengths of geodesics in the rectangular torus $(14, 16)$ using the matrices $\alpha, \beta$ above with $la = 7$ and $lb = 8$ in a similar way that we did in the pair of pants.

## Acknowledgments

These notes concern the results that the author spoke about in a Symposium of the 29th Annual Conference of the Ramanujan Math Society, at IISER, Pune. The author appreciated the kind hospitality at Pune. She also would like to express gratitude to Madhav Modak for his thorough and insightful reading of the manuscript and to Ravi Kulkarni for his encouragement to write it. Also this work was supported by NSF 1105772.

## References

[1] Jayadev Athreya, Alexander Bufetov, Alex Eskin and Maryam Mirzakhani, Lattice point asymptotics and volume growth on teichmüller space, *Duke Mathematical Journal*, **161(6)** (2012) no. 6, 1055–1111.

[2] Ara Basmajian, The stable neighborhood theorem and lengths of closed geodesics, *Proceedings of the American Mathematical Society*, **119** (1993) no. 1, 217–224.

[3] Ara Basmajian, Universal length bounds for non-simple closed geodesics on hyperbolic surfaces, *Journal of Topology* (2013).

[4] Joan S. Birman and Caroline Series, An algorithm for simple curves on surfaces, *J. London Math. Soc. (2)*, **29** (1984) no. 2, 331–342. MR MR744104 (85m:57002)

[5] Joan S. Birman and Caroline Series, Geodesics with bounded intersection number on surfaces are sparsely distributed, *Topology*, 24 (1985) no. 2, 217–225.

[6] Martin Bridson and André Haefliger, *Metric spaces of non-positive curvature*, vol. 319, Springer (1999).

[7] Peter Buser, *Geometry and spectra of compact Riemann surfaces*, Springer (2010).

[8] James W. Cannon, *The growth of the closed surface groups and the compact hyperbolic coxeter groups*, (1980).

[9] James W. Cannon and Philip Wagreich, *Growth functions of surface groups*, Mathematische Annalen 293 (1992) no. 1, 239–257.

[10] Moira Chas, The Goldman bracket and the intersection of curves on surfaces, To appear Contemporary Mathematics, AMS, Proceedings of the conference "Groups, Geometry and Dynamics".

[11] Moira Chas, Combinatorial Lie bialgebras of curves on surfaces, *Topology*, **43** (2004) no. 3, 543–568.

[12] Moira Chas and Steven P. Lalley, Self-intersections in combinatorial topology: statistical structure, *Inventiones Mathematicae*, **188** (2012) no. 2, 429–463.

[13] Moira Chas, Keren Li and Bernard Maskit, Experiments suggesting that the distribution of the hyperbolic length of closed geodesics sampling by word length is Gaussian, To appear in Experimental Mathematics.

[14] Moira Chas and Anthony Phillips, Self-intersection numbers of curves in the doubly-punctured plane, *Experimental Mathematics*, 21 (2012), no. 1, 26–37.

[15] Marshall Cohen and Martin Lustig, Paths of geodesics and geometric intersection numbers. I, Combinatorial group theory and topology (Alta, Utah, 1984), Ann. of Math. Stud., vol. 111, Princeton Univ. Press, Princeton, NJ, (1987), 479–500. MR MR895629 (88m:57016)

[16] Marshall Cohen and Wolfgang Metzler, What does a basis of $F(a,b)$ look like?, *Mathematische Annalen*, **257** (1981) no. 4, 435–445.

[17] William J. Floyd and Steven P. Plotnick, Growth functions on Fuchsian groups and the Euler characteristic, *Inventiones Mathematicae*, **88** (1987) no. 1, 1–29.

[18] Mikhael Gromov, Groups of polynomial growth and expanding maps, *Institut des Hautes 'Etudes Scientifiques. Publications Mathématiques*, (1981) no. 53, 53–73.

[19] Laurent Guillopé, Sur la distribution des longueurs des g'eod'esiques fermées d'une surface compactea bord totalement géodésique, Floer homology, gauge theory, and low-dimensional topology: Proceedings of the Clay Mathematics Institute 2004 Summer School, Alfréd Rényi Institute of Mathematics, Budapest, Hungary, June 5–26, 2004 (1986).
[20] Joel Hass and Peter Scott, Intersections of curves on surfaces, *Israel Journal of Mathematics*, 51 (1985) no. 1, 90–120.
[21] Joel Hass and Peter Scott, Configurations of curves and geodesics on surfaces, Geometry and Topology Monographs (1999).
[22] John Hempel, Traces, lengths, and simplicity for loops on surfaces, *Topology and its Applications*, **18** (1984) no. 2, 153–161.
[23] Michael Hull and Denis Osin, Conjugacy growth of finitely generated groups, *Advances in Mathematics*, **235** (2013) 361–389.
[24] Steven P. Lalley, Renewal theorems in symbolic dynamics, with applications to geodesic flows, noneuclidean tessellations and their fractal limits, *Acta Mathematica*, **163** (1989) no. 1, 1–55.
[25] Steven P. Lalley, Statistical regularities of self-intersection counts for geodesics on negatively curved surfaces, *Duke Mathematical Journal*, **163** (2014) no. 6, 1191–1261.
[26] Martin Lustig, Paths of geodesics and geometric intersection numbers II, Combinatorial group theory and topology (Alta, Utah, 1984), Ann. of Math. Stud., Princeton Univ. Press, Princeton, NJ, vol. 111, (1987) 501–543. MR MR895629 (88m:57016)
[27] John Milnor, A note on curvature and fundamental group, *J. Diff. Geom*, (1968).
[28] Maryam Mirzakhani, Simple geodesics on hyperbolic surfaces and the volume of the moduli space of curves, Ph.D. thesis, Harvard University, (2004).
[29] Maryam Mirzakhani, Growth of the number of simple closed geodesics on hyperbolic surfaces, *Ann. of Math. (2)*, **168** (2008), no. 1, 97–125. MR MR2415399 (2009c:32027)
[30] Lee Mosher, Communications-what is a train track? *Notices of the American Mathematical Society*, **50** (2003) no. 3, 354–356.
[31] Toshihiro Nakanishi *et al.*, The lengths of the closed geodesics on a Riemann surface with self-intersections, *Tohoku Mathematical Journal*, **41** (1989) no. 4, 527–541.
[32] Henri Poincaré, Papers on topology: Analysis situs and its five supplements, *American Mathematical Soc.*, vol. 37, (2010).
[33] Igor Rivin, Simple curves on surfaces, *Geometriae Dedicata*, **87** (2001) no. 1, 345–360.
[34] Igor Rivin, Geodesics with one self-intersection, and other stories, Advances in Mathematics (2012).
[35] Jenya Sapir, Bounds on the number of non-simple geodesics on a surface.
[36] Lower bound for the number of closed geodesics on pairs of pants (and arbitrary surfaces).
[37] Richard Sharp, Periodic orbits of hyperbolic flows, On Some Aspects of the Theory of Anosov Systems, Springer Monographs in Mathematics, Springer Berlin Heidelberg, (2004) 73–138.
[38] Herbert Solomon, *Geometric probability*, SIAM, vol. 28, (1978).
[39] Dennis Sullivan, The density at infinity of a discrete group of hyperbolic motions, *Publications Mathématiques de l'IHÉS*, **50** (1979) no. 1, 171–202.
[40] Mathew Wroten, The eventual gaussian distribution of self-intersection numbers on closed surfaces, Ph.D. Thesis Stony Brook University (2013).
[41] Mathew Wroten, The eventual gaussian distribution for self-intersection numbers on closed surfaces, http://arxiv.org/abs/1405.7951 (2014).
[42] Akira Yamada, On Marden's universal constant of fuchsian groups, *Kodai Mathematical Journal*, **4** (1981) no. 2, 266–277.
[43] Akira Yamada, On Marden's universal constant of fuchsian groups II, *Journal d'Analyse Mathématique*, **41** (1982) no. 1, 234–248.