SELF-INTERSECTIONS OF ARCS ON A PAIR OF PANTS

NHAT MINH DOAN, HANH VO

ABSTRACT. We investigate arcs on a pair of pants and present an algorithm to compute the self-intersection number of an arc. Additionally, we establish bounds for the self-intersection number in terms of the word length. We also prove that the spectrum of self-intersection numbers of 2-low-lying arcs covers all natural numbers.

1. INTRODUCTION

Curves play important roles in the study of surfaces. Dehn studied curves from a topological and combinatorial perspective. He expressed closed curves as finite words in the generators of the fundamental group. He introduced the word problem (characterizing the identity element of the fundamental group), the conjugacy problem (determining conjugate elements), and others. For a long time, mathematicians have been interested in closed curves. The relationship between lengths (combinatorial lengths or hyperbolic lengths) and self-intersection numbers of closed curves has been actively investigated (see [1, 4, 8, 13, 15, 16, 19, 22, 31]). Various algorithms exist to compute self-intersections, such as those found in [2, 10, 17, 18, 27, 29].

We are interested in how the results concerning lengths and self-intersections of closed curves can be reproduced and adapted for *arcs*. Let S be a surface with boundary or punctures, and let ∂S denote the union of these boundary components and punctures. An *arc* is an immersion $\alpha\colon (0,1) \to S^\circ$, $t \mapsto \alpha(t)$, such that each endpoint (as $t \to 0^+$ or $t \to 1^-$) lies in ∂S. Two arcs are considered *equivalent* if they are homotopic relative to their endpoints, with endpoints on boundary components allowed to move along those components, and provided that no arc is homotopic into ∂S. The *self-intersection number* of an arc is the minimum number of transverse self-intersections among all arcs in its equivalence class.

There are various methods to encode the set (or a subset) of arcs on surfaces, such as using continued fractions or left-right cutting sequences as described in [24, 28], or using cutting sequences with sides of an ideal triangulation of the surface, or Farey sequences as detailed in [21]. In this paper, we introduce a different encoding method using *surface word* for arcs that facilitates the computation of the number of self-intersections. We specifically focus on the set of arcs on a pair of pants (a thrice-punctured sphere), a fundamental case that can be extended to more general surfaces with punctures or boundary.

This approach is similar to the technique using the surface word for curves initiated by Chas and Phillips [16]. In particular, the surface word is constructed by choosing a polygonal fundamental domain in the universal cover of the pair of pants, and intersection numbers can then be analyzed by considering how segments of curves or arcs cross this polygon. In the case of arcs, we define the surface word as $a1A3b2B3$, illustrated in Figure 1. Here, 1, 2, and 3 denote the boundary components (or punctures), while a and b represent the seams connecting punctures 1-3 and 2-3, respectively. In this notation, $A = a^{-1}$ and $B = b^{-1}$ are the inverses of a and b.

An arc on a pair of pants corresponds to a word $w = n_1 n_2 \notin \{11, 22\}$ or $w = n_1 x_1 \ldots x_k n_2$ where

- $n_1, n_2 \in \{1, 2, 3\}$,
- $x_i \in \{a, A, b, B\}$ for $1 \le i \le k$,
- $x_i \ne x_{i+1}^{-1}$ for $1 \le i < k$,
- if $n_1 = 1$ then $x_1 \notin \{a, A\}$,
- if $n_1 = 2$ then $x_1 \notin \{b, B\}$,
- if $n_2 = 1$ then $x_k \notin \{a, A\}$,
- if $n_2 = 2$ then $x_k \notin \{b, B\}$.

Remark 1.1. If $n_1 = 3$ or $n_2 = 3$, there are no further restrictions: aside from the two seams joining punctures 1 and 3 or 2 and 3, any arc must pass through the seam between punctures 1 and 2. Also,
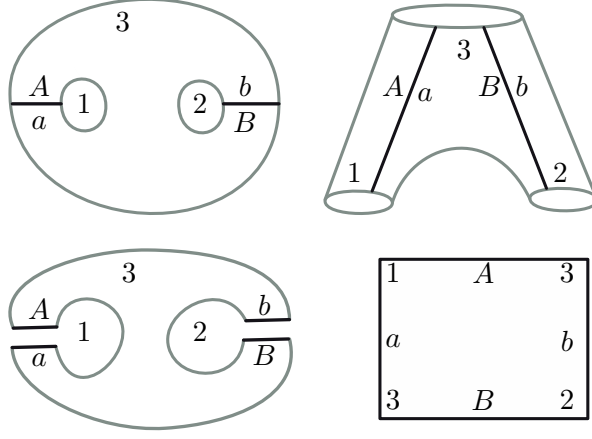
FIGURE 1. The surface word for pairs of pants.

reversing the orientation of an arc always yields a different word except in the case of the arc whose word is 33, where both orientations give the same word.

The word length of w is defined as the number of letters in the word. We denote the word length of w by $\mathrm{WL}(w)$ and the self-intersection number of w by $i(w)$. Our results focus on the topological quantities of arcs, including word lengths, self-intersection numbers and their relations. For a geometric perspective, we refer to [3, 5, 7, 20, 21, 25, 30], which discuss orthogeodesics.

We start by explaining an algorithm to compute the self-intersection number of an arc. This algorithm builds upon techniques developed for computing self-intersection numbers of closed curves on surfaces, specifically motivated from the work of Cohen-Lustig [17] and Despré-Lazarus [18].

Theorem 1.2. *The self-intersection of an arc $w = n_1 x_1 \ldots x_L n_2$ can be computed in $O(L^2)$.*

We also bound the self-intersection number of an arc by its word length.

Theorem 1.3. *Let $w = n_1 x_1 \ldots x_L n_2$. Then $i(w) \leq \frac{L(L+1)}{2}$ and*

$$i(w) \geq \begin{cases} \frac{L}{2} - 1 & \text{if } L \text{ is even} \\ \frac{L-1}{2} & \text{if } L \text{ is odd} \end{cases}$$

The lower bounds are achieved by $1(BA)^{L/2}2$ if L is even and $1(ba)^{(L-1)/2}b1$ if L is odd.

Based on computational experiments (see Table 4), we expect the maximal self-intersection number of w to be $\frac{L^2}{4} + L$ if L is even and to be $\frac{L^2-1}{4} + L$ if L is odd, and they are attained by the self-intersection numbers of the arcs $1(bA)^n 3$ and $3(bA)^n b3$.

Lastly, some recent studies have directed their attention towards counting various types of low-lying curves on the modular surface [6, 9, 12] or in more general contexts [23], as well as exploring intriguing identities of specific families of low-lying arcs on general hyperbolic surfaces [7]. We define an arc as *k-low-lying* if it wraps around the same cuff (a boundary component) consecutively at most k times. Our computations reveal that for any natural number $N \leq 1000$, there exists a 2-low-lying arc whose self-intersection number is N (see Table 11 for $N \leq 89$). Moreover, we prove that this property holds for all $N \in \mathbb{N}$.

Theorem 1.4. *The spectrum of self-intersection numbers of 2-low-lying arcs on a pair of pants covers all natural numbers.*

We remark that low-lying arcs are related to fractions with absolutely bounded entries (or partial quotients) in their continued fractions [28]. Zaremba's conjecture [11, 32] predicts that the denominators of these fractions cover all natural numbers, as detailed in Section 4. The geometric interpretation of the denominator of these fractions is the lambda length of the associated arcs [26]. Our result above addresses a topological aspect of this problem—specifically, the self-intersection numbers of arcs rather than the lambda lengths—and, as we will see in the final section, the proof is straightforward by carefully choosing some special family of low-lying arcs (see Table 10).
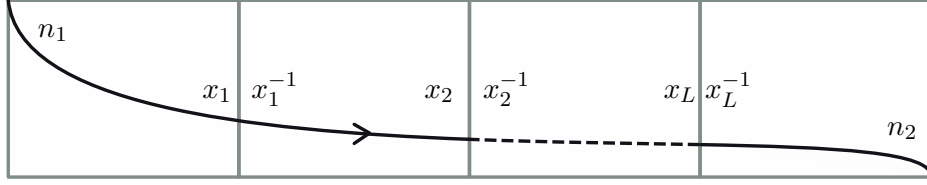
FIGURE 2. A lift of $w = n_1 x_1 \ldots x_L n_2$ to the planar model.

2. COMPUTING SELF-INTERSECTION NUMBERS

Proof of Theorem 1.2. To compute the self-intersection numbers of w, we now describe how to lift it to a planar model for the universal cover induced by the choice of surface word. A lift of w is made of $L + 1$ segments, namely $w_1 = n_1 x_1$, $w_i = x_{i-1}^{-1} x_i$ for $i = 2, \ldots, L$ and $w_{L+1} = x_L n_2$ (see Figure 2). Figure 3 illustrates all possible configurations of a segment w_i in a fundamental domain. A pair (w_i, w_j) is said to be *decidable* if we can decide if there is an intersection point from the current fundamental domain. More precisely, w_i and w_j are decidable if:

- an endpoint of w_i and an endpoint of w_j are in the same corner of the fundamental domain, that is, in a puncture $(1, 2, 3)$, or
- their endpoints are on different edges (a, A, b, B) or corners of the fundamental domain.

It is *undecidable* otherwise, meaning that at least one endpoint of w_i and one endpoint of w_j lie on the same edge of the fundamental domain (a, A, b, B). For example, (aB, Ab) is a decidable pair because their endpoints lie on different edges of the fundamental domain, and (ab, aB) is an undecidable pair. In Figure 4, we illustrate some examples of decidable and undecidable pairs (w_i, w_j): the upper half for decidable and the lower half otherwise. The upper left half of Figure 4 represents an example where there is no intersection, which we shall refer to as a *non-intersecting* decidable pair, while the upper right half represents an example where there is an intersection, which we shall refer to as an *intersecting* decidable one. We notice that in the first case of decidability, namely if an endpoint of w_i and an endpoint of w_j are in the same corner of the fundamental domain, that is, in a puncture $(1, 2, 3)$, the pair (w_i, w_j) is non-intersecting. Note that the order of each segment w_i and w_j is irrelevant for decidability. Thus, the full list of intersecting and non-intersecting decidable pairs (w_i, w_j) (up to order) is presented in Tables 1 and 2.

We first set `intersection` to be 0. For $i = 1, \ldots, L$ and $j = i + 1, \ldots, L + 1$, if the pair (w_i, w_j) is unchecked, we will check if they intersect. In case of decidability, if they intersect, we increase `intersection` by 1 and mark it as checked. In the case of undecidability, namely, at least one endpoint of w_i and one endpoint of w_j are in the same edge of the fundamental domain, that is, a, A, b, B. In this case, we extend w_i (and similar for w_j) by moving to the next segment in the lift of w, namely w_{i+1}, which we shall refer to as *forwarding*, or by moving to the previous segment in the lift of w, namely w_{i-1}, which we shall refer to as *backwarding*. We keep forwarding or backwarding until obtaining a divergent pair (w_i', w_j'), that is, until their other endpoints lie on different edges or corners of the fundamental domain. This process must terminate because the extension must reach the punctures, if not earlier. Note that we mark all the pairs while forwarding or backwarding as checked, namely

$$(w_{i+1}, w_{j+1}), (w_{j+1}, w_{i+1}), \ldots, (w_i', w_j'), (w_j', w_i').$$

After forwarding and backwarding, we reach the configuration where we can decide whether (w_i, w_j) creates an intersection. If so, we increase `intersection` by 1.

Example: In Figure 5, we illustrate an example of extending an undecidable pair w_i, w_j. Denote by `start`(w_i) and `end`(w_i) the starting and ending letter of w_i. In this case `start`$(w_i) = $ `end`(w_j) and
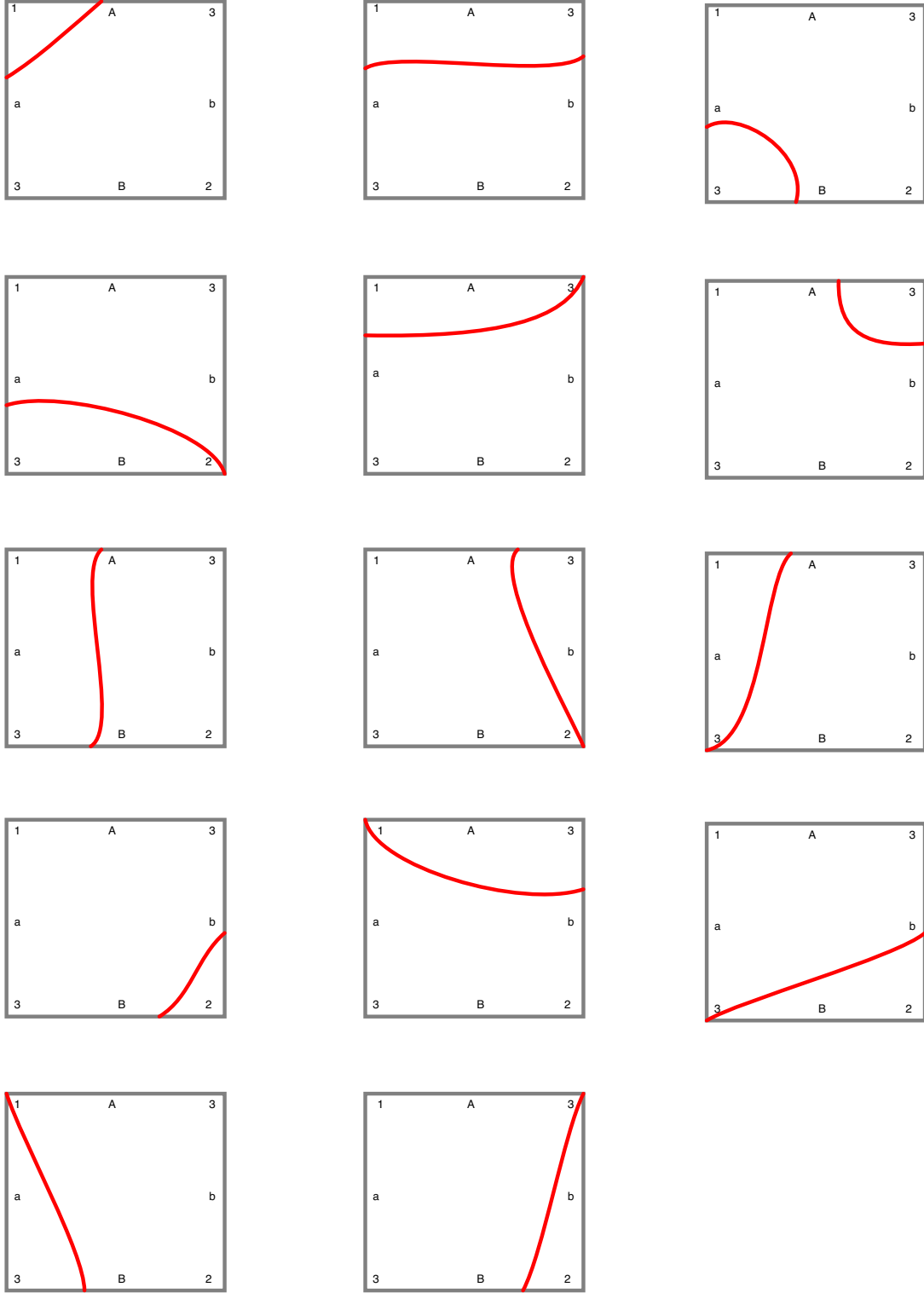
FIGURE 3. All possible configuration of a segment w_i in a fundamental domain.

$\text{end}(w_i) = \text{start}(w_j)$. In one direction, we forward w_i and backward w_j until divergent, we shall call them w_i^1, w_j^1. In the other direction, we forward w_j and backward w_i until divergent, we shall call them w_i^2, w_j^2. There is an intersection in this case because w_j^1 is on the right of w_i^1 whereas w_j^2 is on the left of w_i^2.

The obtained `intersection` number after checking all pairs (w_i, w_j) is the self-intersection number of w. There are $L(L+1)/2$ pairs to check. For each pair we need to extend at most L times, so there are L pairs to check. Thus the self-intersection number of w can be computed in at most $O(L^3)$ steps.
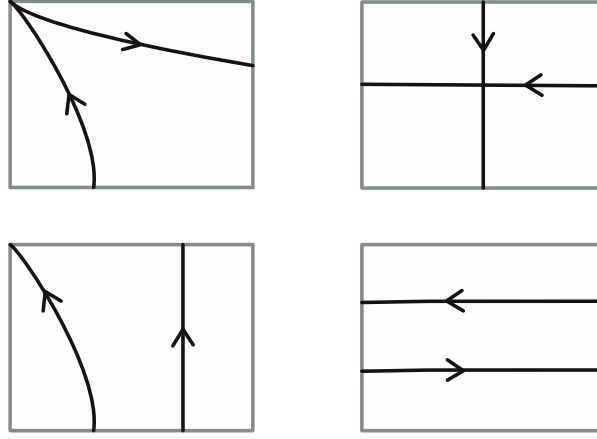
FIGURE 4. Examples: upper half - decidable pairs, lower half - undecidable pairs.

w_i	w_j
aA, Aa	$1b, b1, 1B, B1$
ab, ba	$AB, BA, 2A, A2, 3A, A3, 1B, B1, 3B, B3$
$2a, a2$	$AB, BA, 3A, A3, bB, Bb, 3b, b3, 1B, B1, 3B, B3$
$3a, a3$	$Ab, bA, AB, BA, 2A, A2, 3A, A3, 1b, b1, 1B, B1$
Ab, bA	$3B, B3$
$2A, A2$	$bB, Bb, 1b, b1, 3b, b3, 3B, B3$
$3A, A3$	$aB, Ba, b1, 1b, B1, 1B$
$1b, b1$	$AB, BA, B3, 3B$
$3b, b3$	$aB, Ba, AB, BA, B1, 1B, B3, 3B$

TABLE 1. Intersecting decidable pairs (w_i, w_j).

w_i	w_j
aA, Aa	$bB, Bb, b3, 3b, B3, 3B$
aB, Ba	$Ab, bA, A2, 2A, b1, 1b$
$a2, 2a$	$Ab, bA, A2, 2A, b1, 1b$
$a3, 3a$	$bB, Bb, b3, 3b, B3, 3B$
$A3, 3A$	$B3, 3B$
$B1, 1B$	$Ab, bA, A2, 2A, b1, 1b$

TABLE 2. Non-intersecting decidable pairs (w_i, w_j).

However, as explained in [18][Corollary 19], we only need a quadratic time since each pair of distinct segments is checked at most once. $\qquad\square$

Example 2.1. In this example, we compute the self-intersection number of the arc $w = 1BABA2$. A lift of w is made of 5 segments $w_1 = 1B, w_2 = bA, w_3 = aB, w_4 = bA, w_5 = a2$ (see Figure 6). For $i = 1, \ldots, 4$ and $j = i, \ldots, 5$, we check if a self-intersection of w is coming from the pair (w_i, w_j) (see
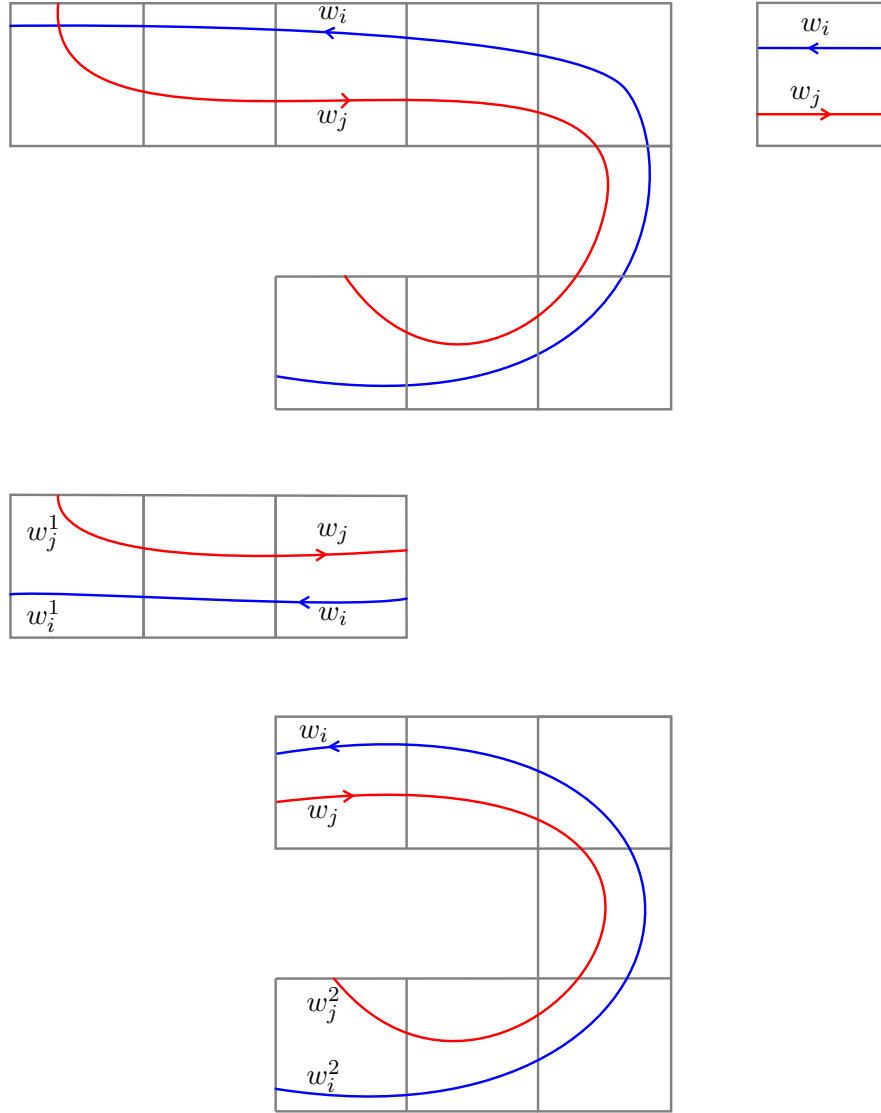
FIGURE 5. Forward and backward w_i, w_j.

Table 3). If i and j have different parity, the pair (w_i, w_j) is a non-intersecting decidable, so they do not contribute intersections. The pair $(w_1, w_3) = (1B, aB)$ is undecidable because both w_1 and w_3 end at B (see Table 3 where X means undecidable). We need to forward this pair to $(w_2, w_4) = (bA, bA)$, which is again undecidable so we continue forward it to $(w_3, w_5) = (aB, a2)$. Now it is decidable: at the beginning, w_3 is on the right of w_1, while at the end of the forwarding process, the corresponding lift of w_3 is on the left of that of w_1. Therefore, there is an intersection point coming from the pair (w_1, w_3) (or equivalently, from the pairs (w_2, w_4) and (w_3, w_5)). We mark all these pairs as checked. Now we only have the pair $(w_1, w_5) = (1B, 2a)$ to check. This pair is an intersecting decidable pair, so it contributes one self-intersection for w. Thus the number of self-intersections of w is 2.

	$w_1 = 1B$	$w_2 = bA$	$w_3 = aB$	$w_4 = bA$	$w_5 = a2$
$w_1 = 1B$		0	X	0	1
$w_2 = bA$			0	X	0
$w_3 = aB$				0	1
$w_4 = bA$					0
$w_5 = a2$					

TABLE 3. Computing the self-intersection number of $1BABA2$, where X means undecidable.
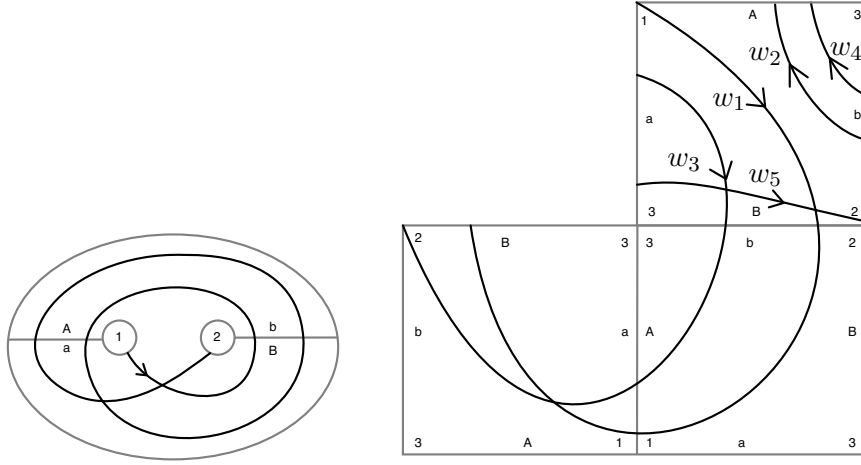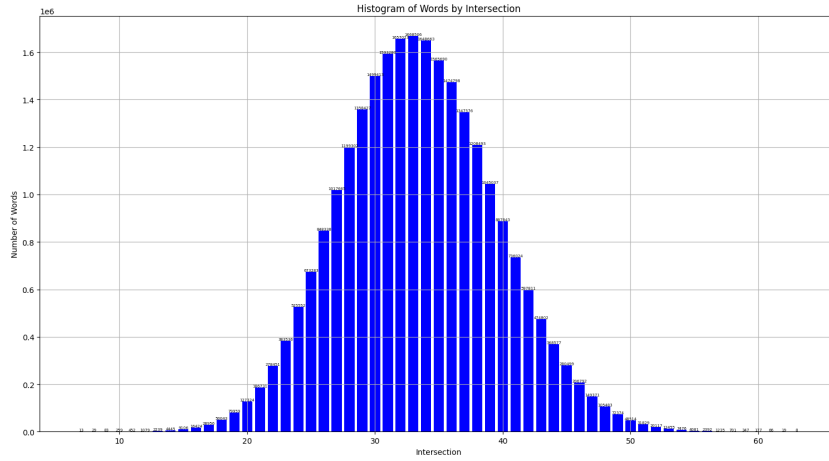
FIGURE 6. Arc $1BABA2$ and the segments $w_1, \ldots, w_5$ of its lifts.

FIGURE 7. Self-intersection number of arcs with word length 16.

Remark 2.2. Our implementation for a pair of pants is available on `https://github.com/hanhv/self-intersections-arcs`. The list of all simple arcs is

$$\{12, 13, 21, 23, 31, 32, 33, 1b1, 1B1, 2a2, 2A2\}.$$

We also computed the self-intersection numbers of all arcs of a given word length and we saw that the distribution of self-intersection numbers is expected to be Gaussian (see Figure 7 for word length is 16). This agrees with the results by Chas-Lalley [14] for closed curves.

Remark 2.3. The idea of the algorithm works for arcs on other surfaces with given surface words. For instance, the surface word of the punctured torus is $ABab$. In addition, the idea also works for computing intersection numbers between two arcs.

3. BOUNDING SELF-INTERSECTION NUMBERS

In this section, we bound the self-intersection number of arcs with a given word length. Lemmas 3.1, 3.2, and Proposition 3.3 adapt results of Chas–Phillips [16][Propositions 4.3, 4.4, and 4.6] from curves to arcs, and the proofs are analogous. Table 4 presents the computational results for the minimum and maximum self-intersection numbers.

A word in the generators of a surface group and their inverses is *positive* if no generator occurs along with its inverse. Note that a positive word is automatically reduced. If w is a word in the alphabet $\{a, A, b, B\}$, we denote by $\alpha(w)$ (resp. $\beta(w)$) the total number of occurrences of a and A (resp. b and B).

Word length	Minimum	Maximum
2	0	0
3	0	1
4	1	3
5	1	5
6	2	8
7	2	11
8	3	15
9	3	19
10	4	24
11	4	29
12	5	35
13	5	41
14	6	48
15	6	55
16	7	63

TABLE 4. Minimum and maximum self-intersection numbers of arcs of given word lengths.

Lemma 3.1. *For any reduced word w in the alphabet a, A, b, B, there is a positive word w' of the same word length with $\alpha(w') = \alpha(w), \beta(w') = \beta(w)$ and $i(w') \leq i(w)$.*

Note that w' is obtained from w by replacing maximal capitalized subwords by ther inverses.

Proof. We can change $w = n_1 x_1 \ldots x_L n_2$ into a word written with $x_i \in \{a, b\}$ for $i = 1, \ldots, L$ while controlling the self-intersection number. If all the letters in x_i are capitals, take $w' = w^{-1}$. Otherwise, look in w for a maximal connected string of (one or more) capital letters. The letters at the ends of this string must be one of the pairs $(A, A), (A, B), (B, A), (B, B)$. In the case (B, B) (the other three cases admit a similar analysis), let that string be $B^{b_1} A^{a_1} B^{b_2} A^{a_2} \ldots B^{b_k}$, and w is either one of the following cases

$$w = n_1 B^{b_1} \ldots B^{b_k} a^{a_k} y, \quad n_1 \neq 2,$$
$$w = x a^{a_0} B^{b_1} \ldots B^{b_k} n_2, \quad n_2 \neq 2,$$
$$w = x a^{a_0} B^{b_1} \ldots B^{b_k} a^{a_k} y,$$

where x, y stand for the rest of the word.

For the first case, consider a representative of w with minimal self-intersection. In this representative, consider the arcs corresponding to the segments $w_1 = n_1 B$ (joining n_1 to the first B of B^{b_1}) and $w_{k+1} = ba$ (joining the last small b that arises from B^{b_k} to the first a in a^{a_k}). These two arcs intersect at a point P. By smoothing the arc at P (see Figure 8), we obtain an arc corresponding to the word

$$w' = n_1 (B^{b_1} \ldots B^{b_k})^{-1} a^{a_k} y, \quad n_1 \neq 2,$$
$$= n_1 b^{b_k} \ldots b^{b_1} a^{a_k} y$$

This word has the same α and β values as w, has lost at least one self-intersection, and has strictly fewer upper-case letters than w.

Similarly, by the smoothing argument, we obtain the arc

$$w' = x a^{a_0} (B^{b_1} \ldots B^{b_k})^{-1} n_2, \quad n_2 \neq 2,$$

for the second case, and

$$w' = x a^{a_0} (B^{b_1} \ldots B^{b_k})^{-1} a^{a_k} y$$

for the third case. In either case, w' has the same α and β values as w, has lost at least one self-intersection, and has strictly fewer upper-case letters than w. The process may be repeated until all upper-case letters have been eliminated. $\square$
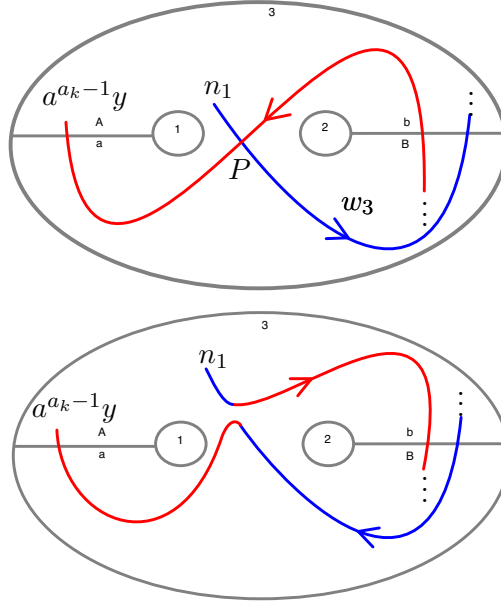
FIGURE 8. Smoothing the arc w at P.

Lemma 3.2. *Let w be a non-simple arc. Then there exist an arc u and a curve v such that the word for w can be written as the insertion of the word for v into the word for u and*

$$i(u) + i(v) + 1 \leq i(w).$$

Proof. Consider a minimal representative of w. It must have self-intersections. Write $w = x_1 \ldots x_L$, where $x_1, x_L \in \{1, 2, 3\}$ and $x_2, \ldots, x_{L-1} \in \{a, A, b, B\}$. (This is the only time we write $w = x_1 \ldots x_L$ instead of $n_1 x_1 \ldots x_L n_2$.) Suppose that $x_i x_{i+1}$ and $x_j x_{j+1}$, with $i < j$, are the two segments intersecting at P. By smoothing this intersection point, we obtain the arc $u = x_1 \ldots x_i x_{j+1} \ldots x_L$ and the curve $v = x_{i+1} \ldots x_j$ as in the statement of the lemma. (In case $i + 1 = j$, v is a single-letter word.) The change from w to $u \cup v$ does not add any new intersections, while the intersection corresponding to P is lost. Hence $i(u) + i(v) + 1 \leq i(w)$. $\qquad\square$

The following proposition is a result for positive words representing closed curves.

Proposition 3.3. *If w is a positive word representing an arc then $i(w) \geq \alpha(w) - 1$ and $i(w) \geq \beta(w) - 1$.*

Proof. If the word length of w is at most 3, then both $\alpha(w)$ and $\beta(w)$ are at most 1, hence the statement is true. If word length w is $L \geq 4$, then w is non-simple. Suppose that the statement holds for all positive words of word length less than L. By Lemma 3.2, w can be written as the concatenation $w = u \cdot v$ of an arc u and a curve v satisfy $i(u) + i(v) + 1 \leq i(w)$. Note that both u and v are positive words and of word length less than L. By the induction hypothesis, $i(u) \geq \alpha(u)$ and $i(u) \geq \beta(u)$. Using [16][Proposition 4.6] for the curve v, $i(v) \geq \alpha(v)$ and $i(v) \geq \beta(v)$. Therefore,

$$i(u) + i(v) \geq \alpha(u) - 1 + \alpha(v) - 1.$$

Hence,

$$i(w) \geq i(u) + i(v) + 1 = \alpha(u) + \alpha(v) - 1 = \alpha(w) - 1.$$

The β inequality is proved in the same way. $\qquad\square$

Proof of Theorem 1.3. By Lemma 3.1, there is a positive word w' of the same word length as w such that $\alpha(w') = \alpha(w), \beta(w') = \beta(w)$, and $i(w') \leq i(w)$. By Proposition 3.3,

$$i(w') \geq \max\{\alpha(w) - 1, \beta(w) - 1\}.$$

Since $\alpha(w) + \beta(w) = L$, it follows that

$$i(w) \geq \frac{L}{2} - 1$$

if L is even and

$$i(w) \geq \frac{L+1}{2} - 1 = \frac{L-1}{2}$$

if L is odd. $\qquad\square$

We now show that the lower bounds are attainable.

Lemma 3.4. *For all $n \in \mathbb{N}$,*

$$i(1(BA)^n 2) = n, \quad i(1(bA)^n 3) = n^2 + 2n,$$

$$i(1(BA)^n B1) = n, \quad i(3(bA)^n b3) = n^2 + 3n + 1.$$

Proof. We first compute $i(1(BA)^n 2)$. We need to check all pair (w_i, w_j) for $i = 1, \ldots, 2n$ and $j = i, \ldots, 2n+1$. We have $w_1 = 1B, w_{2n+1} = a2$, $w_i = bA$ for even i and $w_i = aB$ for odd $i \geq 3$. If i and j have different parity, the pair (w_i, w_j) is a non-intersecting decidable. For odd $i \geq 3$, the pair $(w_1, w_i) = (1B, aB)$ is undecidable. We forward them until reaching $(w_., w_{2n+1}) = (aB, a2)$, obtaining one intersection point. Note that while forwarding them, we mark all pairs $(w_e, w_{e'}) = (bA, bA)$ for even e, e' and $(w_o, w_{o'}) = (aB, aB)$ for odd $o, o' (o \geq 3)$ (see Table 5). In total, we have n intersections that come from pairs (w_1, w_o) for odd $3 \leq o \leq 2n+1$. Thus, the self-intersection numbers of $1(BA)^n 2$ is

$$i(1(BA)^n 2) = \frac{(2n-1)-1}{2} + 1 = n.$$

	w_1 $= 1B$	w_2 $= bA$	w_3 $= aB$	w_4 $= bA$	w_{2n-1} $= aB$	w_{2n} $= bA$	w_{2n+1} $= a2$
$w_1 = 1B$		0	X	0	X	0	1
$w_2 = bA$			0	X	0	X	0
$w_3 = aB$				0	X	0	1
$w_4 = bA$					0	X	0
$w_{2n-1} = aB$						0	1
$w_{2n} = bA$							0
$w_{2n+1} = a2$							

TABLE 5. Computing the self-intersection number of $1(BA)^n 2$.

We do similarly to compute the self-intersection numbers of other arcs. Table 6 is for $1(bA)^n 3$: its self-intersection number is

$$i(1(bA)^n 3) = n \times 1 + 2 \times 1 + 2 \times 2 + \cdots + 2 \times n = n^2 + 2n.$$

	w_1 $= 1b$	w_2 $= BA$	w_3 $= ab$	w_4 $= BA$	w_{2n-1} $= ab$	w_{2n} $= BA$	w_{2n+1} $= a3$
$w_1 = 1b$		1	X	1	X	1	1
$w_2 = BA$			1	X	1	X	1
$w_3 = ab$				1	X	1	1
$w_4 = BA$					1	X	1
$w_{2n-1} = ab$						1	1
$w_{2n} = BA$							1
$w_{2n+1} = a3$							

TABLE 6. Computing the self-intersection number of $1(bA)^n 3$.

Table 7 is for $1(BA)^n B1$: its self-intersection number is

$$i(1(BA)^n B1) = \frac{2n-2}{2} + 1 = n.$$

Table 8 is for $3(bA)^n b3$: its self-intersection number is

$$i(3(bA)^n b3) = n \times 1 + 1 + 3 + 5 + 7 + \cdots + (2n+1) = n + (n+1)^2 = n^2 + 3n + 1.$$

$\square$

	w_1 $= 1B$	w_2 $= bA$	w_3 $= aB$	w_4 $= bA$	w_{2n-1} $= aB$	w_{2n} $= bA$	w_{2n+1} $= aB$	w_{2n+2} $= b1$
$w_1 = 1B$		0	X	0	X	0	X	0
$w_2 = bA$			0	X	0	X	0	1
$w_3 = aB$				0	X	0	X	0
$w_4 = bA$					0	X	0	1
$w_{2n-1} = aB$						0	X	0
$w_{2n} = bA$							0	1
$w_{2n+1} = aB$								0
$w_{2n+2} = b1$								

TABLE 7. Computing the self-intersection number of $1(BA)^n B1$

	w_1 $= 3b$	w_2 $= BA$	w_3 $= ab$	w_4 $= BA$	w_{2n-1} $= ab$	w_{2n} $= BA$	w_{2n+1} $= ab$	w_{2n+2} $= B3$
$w_1 = 3b$		1	X	1	X	1	X	1
$w_2 = BA$			1	X	1	X	1	1
$w_3 = ab$				1	X	1	X	1
$w_4 = BA$					1	X	1	1
$w_{2n-1} = ab$						1	X	1
$w_{2n} = BA$							1	1
$w_{2n+1} = ab$								1
$w_{2n+2} = B3$								

TABLE 8. Computing the self-intersection number of $3(bA)^n b3$.

Remark 3.5. We also attempted to obtain a sharp upper bound for the self-intersections of $w = n_1 x_1 \ldots x_L n_2$ using the method in [16, Section 2.2]. However, that method does not apply because we lack cyclic permutations for arcs. Therefore, our current upper bound is derived from Theorem 1.2. The self-intersection number of w is bounded from above by the total number of pairs (w_i, w_j) for $i = 1, \ldots, L$ and $j = i, \ldots, L + 1$, hence

$$i(w) \leq \frac{L(L+1)}{2} - U(w),$$

where $U(w)$ is the number of undecidable pairs. Based on computational experiments (see Table 4), we expect the maximal self-intersection number of w to be $\frac{L^2}{4} + L$ if L is even and to be $\frac{L^2-1}{4} + L$ if L is odd, and they are attained by the self-intersection numbers of the arcs $1(bA)^n 3$ and $3(bA)^n b3$ as in Lemma 3.4.

4. LOW-LYING ARCS

We conclude this note with a discussion on the self-intersection numbers of low-lying arcs and the proof of Theorem 1.4. An arc is called *k-low-lying* if it wraps around the same cuff (a boundary component) consecutively at most k times. In terms of words, for example, if the arc $w = n_1 x_1 \ldots x_L n_2$ satisfies the condition that for all $i \leq L$,

$$x_i x_{i+1} x_{i+2} x_{i+3} \notin \{aaaa, AAAA, bbbb, BBBB\} \quad \text{if } i + 3 \leq L, L \geq 4,$$

and

$$x_i x_{i+1} \ldots x_{i+6} x_{i+7} \notin \{abababab, ABABABAB, babababa, BABABABA\} \quad \text{if } i + 7 \leq L, L \geq 8,$$

then w is 4-low-lying.

Our computations reveal that for any natural number $N \leq 1000$, there exists a low-lying arc whose self-intersection number is N. Theorem 1.4 extends this result to all natural numbers $N \in \mathbb{N}$, providing a comprehensive view of the behavior of low-lying arcs in this context.

Remark 4.1. The observation about self-intersection numbers of low-lying arcs is somewhat analogous to Zaremba's conjecture [11, 32], which predicts the existence of some natural number $A > 1$ such

that

$$D_A := \left\{ d \in \mathbb{N} : \exists (b, d) = 1 \text{ with } \frac{b}{d} \in R_A \right\} = \mathbb{N}$$

where

$$R_A := \left\{ \frac{b}{d} = [a_1, a_2, ..., a_k] : 0 < b < d, (b, d) = 1 \text{ and } \forall i, a_i \leq A \right\},$$

and

$$[a_1, a_2, ..., a_k] := \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_k}}}}.$$

In our case, geometrically, a fraction $\frac{b}{d}$ corresponds to a geodesic arc on a hyperbolic pair of pants with three cusps $\mathscr{P} = \mathbb{H}/\Gamma(2)$ (see Figure 9), where $\mathbb{H}$ denotes the Poincaré upper half plane, and

$$\Gamma(2) = \left\langle \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \right\rangle.$$
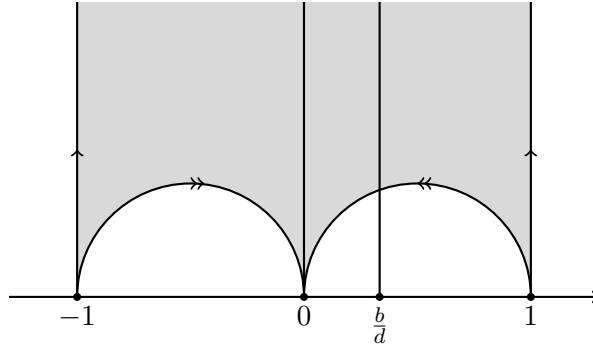

FIGURE 9. A fundamental domain for the hyperbolic pair of pants $\mathscr{P}$ of three cusps, and a lift of the orthogeodesic $\mu_{\frac{b}{d}}$.

The set of lifts of cusps of $\mathscr{P}$ is the union of the orbits of cusps $0, 1$ and ∞

$$\Gamma(2) \cdot 0 \cup \Gamma(2) \cdot 1 \cup \Gamma(2) \cdot \infty = \mathbb{Q} \cup \{\infty\}.$$

An orthogeodesic on $\mathscr{P}$ is a geometric realization of an arc on $\mathscr{P}$. Due to the symmetry of $\mathscr{P}$, we will consider only the set of orthogeodesics lifting to vertical lines from ∞ to fractions between 0 and 1 on the universal cover of $\mathscr{P}$ (identified to $\mathbb{H}$). We denote by $\mu_{\frac{b}{d}}$ the orthogeodesic on $\mathscr{P}$ lifting to the vertical line from ∞ to $\frac{b}{d} \in \mathbb{Q} \cap [0, 1]$. Theorem A in [28] implies that if $\frac{b}{d} = [a_1, a_2, \ldots, a_k]$, then the left-right cutting sequence of $\mu_{\frac{b}{d}}$ is $R^{a_1 - 1}L^{a_2} \ldots X^{a_k - 1}$, where $X = L$ (or R) if k is even (or odd). Thus the "low" in the values of a_i implies the "low" in the wrapping numbers of the orthogeodesic $\mu_{\frac{b}{d}}$ around each cuff consecutively. In these terms, the question about self-intersections of low-lying arcs can be stated as follows:

> Is there a natural number $A > 1$ such that for every $N \in \mathbb{N}$, there exists a fraction $\frac{b}{d} \in R_A$ for which the self-intersection number of $\mu_{\frac{b}{d}}$ equals N?

Theorem 1.4 gives a positive answer to this question for $A = 2$. In the following lemma, the subset $A_i \subset \mathbb{N}$ represents numbers that are realized as self-intersection numbers for a family of arcs which are made explicit in Table 10.

Lemma 4.2. *Let*

$$A_1 := \{(m + n + 1)^2 + 2m + n \,|\, m, n \in \mathbb{N}, m \geq 1\}, \quad A_2 := \{(m + n)^2 + 2m + 3n \,|\, m, n \in \mathbb{N}, m \geq 1\},$$

$$A_3 := \{(n + 4)^2 - 2 \,|\, n \in \mathbb{N}\}, \quad A_4 := \{n(n + 3) \,|\, n \in \mathbb{N}\}, \quad A_5 := \{n(n + 3) + 1 \,|\, n \in \mathbb{N}\},$$

then

$$\mathbb{N} = \{2, 7\} \cup \bigcup_{i=1}^{5} A_i.$$

Proof. We observe that:

$$A_1 = \{(m+n+2)^2 - (n+3) | m, n \in \mathbb{N}, m \geq 1\} = \bigcup_{j=4}^{\infty} \left(\bigcup_{i=-j}^{-3} \{j^2 + i\} \right) \cup \{6\},$$

$$A_2 = \{(m+n+1)^2 + (n-1) | m, n \in \mathbb{N}, m \geq 1\} = \bigcup_{j=4}^{\infty} \left(\bigcup_{i=-1}^{j-3} \{j^2 + i\} \right) \cup \{3, 8, 9\},$$

$$A_3 = \{(n+4)^2 - 2 | n \in \mathbb{N}\} = \bigcup_{j=4}^{\infty} \{j^2 - 2\},$$

$$A_4 = \{(n+1)^2 + (n+1) - 2 | n \in \mathbb{N}\} = \bigcup_{j=4}^{\infty} \{j^2 + j - 2\} \cup \{0, 4, 10\},$$

$$A_5 = \{(n+1)^2 + (n+1) - 1 | n \in \mathbb{N}\} = \bigcup_{j=4}^{\infty} \{j^2 + j - 1\} \cup \{1, 5, 11\},$$

Thus

$$\bigcup_{i=1}^{5} A_i = \bigcup_{j=4}^{\infty} \left(\bigcup_{i=-j}^{j-1} \{j^2 + i\} \right) \cup \{0, 1, 3, 4, 5, 6, 8, 9, 10, 11\} = \mathbb{N} - \{2, 7\}.$$

$\square$

Proof of Theorem 1.4. Consider the low-lying arcs $w = 1(bA)^n(baa)^m 2$. Table 9 presents how we count its self-intersection numbers, with

$$w_1 = 1b, w_2 = BA, w_3 = ab, w_4 = BA, \ldots, w_{2n-1} = ab, w_{2n} = BA, w_{2n+1} = ab,$$

$w'_1 := w_{2n+2} = Ba, w'_2 := w_{2n+3} = Aa, w'_3 := w_{2n+4} = Ab, \ldots, w'_{3m-2} = Ba, w'_{3m-1} = Aa, w'_{3m} = A2$.
Applying the algorithm from Theorem 1.2, it is straightforward to check which segments from w are decidable (see Table 9). For the undecidable pairs, after applying forward and backward operations, we reach a configuration where we can determine whether a pair creates an intersection. Therefore the self-intersection number is

$$(2n + 2(n-1) + \cdots + 2 \times 1) + (2m + 2(m-1) + \cdots + 2 \times 1 - 1) + (n(2m+1) + n + (m+1))$$

$$= (n + m)^2 + 3n + 2m.$$

	w_1	w_2	w_3	w_4	w_{2n-1}	w_{2n}	w_{2n+1}	w'_1	w'_2	w'_3	w'_4	w'_5	w'_6	w'_{3m-2}	w'_{3m-1}	w'_{3m}
w_1		1	X	1	X	1	X	0	1	X	0	1	X	0	1	1
w_2			1	X	1	X	1	1	X	X	0	X	X	0	X	X
w_3				1	X	1	X	1	1	X	1	1	X	1	1	1
w_4					1	X	1	1	X	X	0	X	X	0	X	X
w_{2n-1}						1	X	1	1	X	1	1	X	1	1	1
w_{2n}							1	1	X	X	0	X	X	0	X	X
w_{2n+1}								1	1	X	1	1	X	1	1	1
w'_1									X	0	X	X	0	X	X	0
w'_2										1	X	X	1	X	X	1
w'_3											0	1	X	0	1	1
w'_4												X	0	X	X	0
w'_5													1	X	X	1
w'_6														0	1	1
w'_{3m-2}															X	0
w'_{3m-1}																1
w'_{3m}																

TABLE 9. Computing the self-intersection number of $1(bA)^n(baa)^m 2$, where X means undecidable.

Similarly, we can compute the self-intersection numbers of the other low-lying arcs as in Table 10.

Low-lying arcs w	$i(w)$	Conditions	Continued fraction of $\frac{b}{d}$
$1(bA)^n bab(ABA)^m 2$	$(m+n+1)^2 + 2m+n$	$\forall m,n \in \mathbb{N}, m \geq 1$	$[2,\dots,2,1,2,1,1,\ \underbrace{2,\dots,2}_{2m-1 \text{ times}},1]$ (with $\underbrace{2,\dots,2}_{2n \text{ times}}$)
$1(bA)^n (baa)^m 2$	$(m+n)^2 + 2m+3n$	$\forall m,n \in \mathbb{N}, m \geq 1$	$[\underbrace{2,\dots,2}_{2n \text{ times}},1,1,\underbrace{2,\dots,2}_{2m-1 \text{ times}},1]$
$1bABabaBA(bA)^n 3$	$(n+4)^2 - 2$	$\forall n \in \mathbb{N}$	$[2,1,1,1,1,2,1,1,1,1,\underbrace{2,\dots,2}_{2n \text{ times}},1]$
$1(bA)^n b1$	$n(n+3)$	$\forall n \in \mathbb{N}$	$[\underbrace{2,\dots,2}_{2n \text{ times}},1,1]$
$1(bA)^n ba2$	$n(n+3)+1$	$\forall n \in \mathbb{N}$	$[\underbrace{2,\dots,2}_{2n \text{ times}},1,1,1]$
$1bA2$	2		$[2,1,1]$
$1bAba3$	7		$[2,2,1,1,1,1]$

TABLE 10. Some low-lying arcs and their corresponding self-intersection numbers.

By Lemma 4.2, the spectrum of self-intersection numbers of the 2-low-lying arcs in Table 10 covers all natural numbers. □

TABLE 11. Examples of some low-lying words and their self-intersection numbers.

w	$i(w)$	w	$i(w)$
31	0	3a3	1
3ab1	2	3aB1	3
3aba3	4	3aBa3	5
3abAb3	6	3abAb1	7
3aBaB1	8	3abAbA2	9
3AbAba3	10	3AbAbA3	11
3abAbaB1	12	3abABaB1	13
3ABaBaB1	14	3bAbAbA2	15
3abAbaBa3	16	3abAbAbA3	17
3AbAbAba3	18	3AbAbAbA3	19
3abAbAbAb3	20	3aBabAbAb1	21
3aBaBaBABb1	22	3ABaBaBaB1	23
3bAbAbAbA2	24	3abAbAbAbA2	25
3aBabAbAba3	26	3aBaBaBaBA3	27
3AbAbAbAba3	28	3AbAbAbAbA3	29
3abAbAbaBaB1	30	3abAbABaBaB1	31
3aBabAbAbAb1	32	3aBaBaBaBAb1	33
3ABaBaBaBaB1	34	3bAbAbAbAbA2	35
3abAbAbaBaBa3	36	3abAbABaBaBa3	37
3aBabAbAbAba3	38	3aBaBaBaBaBA3	39
3AbAbAbAbAba3	40	3AbAbAbAbAbA3	41
3abAbAbAbAbAb3	42	3abABabAbAbAb1	43
3aBabABaBaBaB1	44	3AbABaBaBaBaB1	45
3ABabAbAbAbAb1	46	3ABaBaBaBaBaB1	47
3bAbAbAbAbAbA2	48	3abAbAbAbAbAbA2	49
3abABabAbAbAba3	50	3aBabAbAbAbAbA3	51
3AbAbAbAbAbaBa3	52	3ABaBaBaBaBaBA3	53
3baBaBaBaBaBab3	54	3bAbAbAbAbAbAb3	55
3abAbAbAbaBaBaB1	56	3abAbAbABaBaBaB1	57
3abABabAbAbAbAb1	58	3aBabABaBaBaBaB1	59
3AbABaBaBaBaBaB1	60	3ABabAbAbAbAbAb1	61
3ABaBaBaBaBaBaB1	62	3bAbAbAbAbAbAbA2	63
3abAbAbAbaBaBaBa3	64	3abAbAbABaBaBaBa3	65
3abABabAbAbAbAba3	66	3aBabAbAbAbAbAbA3	67

w	$i(w)$	w	$i(w)$
3AbAbAbAbAbAbaBa3	68	3ABaBaBaBaBaBaBA3	69
3baBaBaBaBaBaBaBab3	70	3bAbAbAbAbAbAbAb3	71
3abAbAbAbAbAbAbAb3	72	3abAbABabAbAbAbAb1	73
3abABabABaBaBaBaB1	74	3aBabAbAbAbAbAbaB1	75
3aBabABaBaBaBaBaB1	76	3AbABaBaBaBaBaBaB1	77
3ABabAbAbAbAbAb1	78	3ABaBaBaBaBaBaBaB1	79
3bAbAbAbAbAbAbAbA2	80	3abAbAbAbAbAbAbAbA2	81
3abAbABabAbAbAbAba3	82	3abABabABaBaBaBaBa3	83
3abABaBaBaBaBaBaBA3	84	3aBabAbAbAbAbAbAbA3	85
3AbAbAbAbAbAbaBa3	86	3ABaBaBaBaBaBaBaBA3	87
3baBaBaBaBaBaBaBab3	88	3bAbAbAbAbAbAbAbAb3	89

References

[1] Macarena Arenas and Max Neumann-Coto. Taut smoothings of arcs and curves. *arXiv preprint arXiv:2402.06623*, 2024.

[2] Chris Arettines. A combinatorial algorithm for visualizing representatives with minimal self-intersection. *J. Knot Theory Ramifications*, 24(11):1550058, 17, 2015.

[3] Ara Basmajian. The orthogonal spectrum of a hyperbolic manifold. *American Journal of Mathematics*, 115(5):1139–1159, 1993.

[4] Ara Basmajian. Universal length bounds for non-simple closed geodesics on hyperbolic surfaces. *J. Topol.*, 6(2):513–524, 2013.

[5] Ara Basmajian and Federica Fanoni. Orthosystoles and orthokissing numbers. *arXiv preprint arXiv:2405.01732*, 2024.

[6] Ara Basmajian and Mingkun Liu. Low-lying geodesics on the modular surface and necklaces. *arXiv preprint arXiv:2311.16041*, 2023.

[7] Ara Basmajian, Hugo Parlier, and Ser Peow Tan. Prime orthogeodesics, concave cores and families of identities on hyperbolic surfaces. *Advances in Mathematics*, 460:110026, 2025.

[8] Ara Basmajian, Hugo Parlier, and Hanh Vo. The shortest non-simple closed geodesics on hyperbolic surfaces. *Mathematische Zeitschrift*, 306(1):8, 2024.

[9] Ara Basmajian and Robert Suzzi Valli. Combinatorial growth in the modular group. *Groups, Geometry, and Dynamics*, 16(2):683–703, 2022.

[10] Joan S Birman and Caroline Series. An algorithm for simple curves on surfaces. *Journal of the London Mathematical Society*, 2(2):331–342, 1984.

[11] Jean Bourgain and Alex Kontorovich. On Zaremba's conjecture. *Annals of Mathematics*, pages 137–196, 2014.

[12] Jean Bourgain and Alex Kontorovich. Beyond expansion ii: low-lying fundamental geodesics. *Journal of the European Mathematical Society*, 19(5):1331–1359, 2017.

[13] Moira Chas. Relations between word length, hyperbolic length and self-intersection number of curves on surfaces. In *Recent advances in mathematics*, volume 21 of *Ramanujan Math. Soc. Lect. Notes Ser.*, pages 45–75. Ramanujan Math. Soc., Mysore, 2015.

[14] Moira Chas and Steven P. Lalley. Self-intersections in combinatorial topology: statistical structure. *Invent. Math.*, 188(2):429–463, 2012.

[15] Moira Chas and Anthony Phillips. Self-intersection numbers of curves on the punctured torus. *Experiment. Math.*, 19(2):129–148, 2010.

[16] Moira Chas and Anthony Phillips. Self-intersection numbers of curves in the doubly punctured plane. *Exp. Math.*, 21(1):26–37, 2012.

[17] Marshall Cohen and Martin Lustig. Paths of geodesics and geometric intersection numbers. I. In *Combinatorial group theory and topology (Alta, Utah, 1984)*, volume 111 of *Ann. of Math. Stud.*, pages 479–500. Princeton Univ. Press, Princeton, NJ, 1987.

[18] Vincent Despré and Francis Lazarus. Computing the geometric intersection number of curves. *J. ACM*, 66(6):Art. 45, 49, 2019.

[19] El Hadji Abdou Aziz Diop and Masseye Gaye. Self-intersection on pair of pants. In *Nonlinear Analysis, Geometry and Applications: Proceedings of the Second NLAGA-BIRS Symposium, Cap Skirring, Senegal, January 25–30, 2022*, pages 429–454. Springer, 2022.

[20] Nhat Minh Doan. Geometric filling curves on punctured surfaces. *Glasgow Mathematical Journal*, 65(2):383–400, 2023.

[21] Nhat Minh Doan. Ortho-integral surfaces. *Advances in Mathematics*, 466:110162, 2025.

[22] Viveka Erlandsson and Hugo Parlier. Short closed geodesics with self-intersections. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 169, pages 623–638. Cambridge University Press, 2020.

[23] Viveka Erlandsson and Juan Souto. Counting and equidistribution of reciprocal geodesics and dihedral groups. *International Mathematics Research Notices*, page rnad156, 2024.

[24] François Labourie and Ser Peow Tan. The probabilistic nature of McShane's identity: planar tree coding of simple loops. *Geometriae Dedicata*, 192:245–266, 2018.

[25] Hugo Parlier. Geodesic and orthogeodesic identities on hyperbolic surfaces. *arXiv preprint arXiv:2004.09078*, 2020.

[26] Robert C Penner. The decorated Teichmüller space of punctured surfaces. *Communications in Mathematical Physics*, 113:299–339, 1987.

[27] James Rickards. Computing intersections of closed geodesics on the modular curve. *Journal of Number Theory*, 225:374–408, 2021.

[28] Caroline Series. The modular surface and continued fractions. *Journal of the London Mathematical Society*, 2(1):69–80, 1985.

[29] Ser Peow Tan. Self-intersections of curves on surfaces. *Geom. Dedicata*, 62(2):209–225, 1996.

[30] Marie Trin. Counting arcs of the same type. *arXiv preprint arXiv:2306.07573*, 2023.

[31] Hanh Vo. Short closed geodesics on cusped hyperbolic surfaces. *Pacific Journal of Mathematics*, 318(1):127–151, 2022.

[32] Stanisław K Zaremba. La méthode des "bons treillis" pour le calcul des intégrales multiples. In *Applications of number theory to numerical analysis*, pages 39–119. Elsevier, 1972.

Department of Mathematics, National University of Singapore, Singapore
& Institute of Mathematics, Vietnam Academy of Science and Technology, Hanoi, Vietnam
  Email address: minh.dn@nus.edu.sg, dnminh@math.ac.vn
  URL: https://sites.google.com/view/dnminh

School of Mathematical and Statistical Sciences, Arizona State University
  Email address: hanhmfa@gmail.com
  URL: https://hanhv.github.io