

## Minimal pseudo-Anosov translation lengths on the complex of curves

VAIBHAV GADRE  
CHIA-YEN TSAI

We establish bounds on the minimal asymptotic pseudo-Anosov translation lengths on the complex of curves of orientable surfaces. In particular, for a closed surface with genus  $g \geq 2$ , we show that there are positive constants  $a_1 < a_2$  such that the minimal translation length is bounded below and above by  $a_1/g^2$  and  $a_2/g^2$ .

30F60, 32G15

### 1 Introduction

Let  $S_{g,n}$  be an orientable surface with genus  $g$  and  $n$  punctures. For simplicity, we shall drop the subscripts and denote it by  $S$ . The complex of curves  $\mathcal{C}(S)$  is a locally infinite simplicial complex whose vertices are the isotopy classes of essential, nonperipheral, simple closed curves on  $S$ . A collection of vertices span a simplex if the curves can be isotoped to be disjoint or minimally intersecting on  $S$ . Here, we will assume that the surface  $S$  is nonsporadic, ie, the complexity  $\xi(S) = 3g - 3 + n \geq 2$ . For sporadic surfaces, the complex of curves  $\mathcal{C}(S)$  is either trivial or well-understood.

The mapping class group  $\text{Mod}(S)$  is the group of isotopy classes of diffeomorphisms of  $S$ . This group acts on  $\mathcal{C}(S)$  in the obvious way. Thurston classified the elements of  $\text{Mod}(S)$  into three types: finite order, reducible or pseudo-Anosov. Given a mapping class  $f \in \text{Mod}(S)$ , its asymptotic translation length on  $\mathcal{C}(S)$  is defined to be

$$\ell_{\mathcal{C}}(f) = \liminf_{j \rightarrow \infty} \frac{d_{\mathcal{C}}(\alpha, f^j(\alpha))}{j}$$

where  $\alpha$  is a simple closed curve on  $S$ . The above limit remains unchanged when the numerator is changed by an additive constant. Hence, by the triangle inequality for  $d_{\mathcal{C}}$ , the quantity  $\ell_{\mathcal{C}}(f)$  is independent of the choice of curve  $\alpha$ .

In [5], Masur and Minsky proved that  $f \in \text{Mod}(S)$  is pseudo-Anosov if and only if  $\ell_{\mathcal{C}}(f) > 0$ . In [2], Bowditch refined this, proving that the set of translation lengths of

pseudo-Anosov elements is a subset of  $\mathbb{Q}_+$  with bounded denominators. We denote the minimal positive number in this set by

$$L_C(\text{Mod}(S)) = \min\{\ell_C(f) \mid f \in \text{Mod}(S), \text{ pseudo-Anosov}\}$$

For closed surfaces, Farb, Leininger and Margalit [3] proved that when  $g \geq 2$ ,

$$L_C(\text{Mod}(S)) < \frac{4 \log(2 + \sqrt{3})}{g \log(g - 1/2)}.$$

Here, we find a better upper bound for  $L_C(\text{Mod}(S))$ . Moreover, we also show that a lower bound of the same order holds. To be precise, we show:

**Theorem 1.1** *For closed surfaces with  $g \geq 2$ ,*

$$\frac{1}{162(2g - 2)^2 + 30(2g - 2)} < L_C(\text{Mod}(S)) \leq \frac{4}{g^2 + g - 4}.$$

The upper bound is established by bounding  $\ell_C(f)$  in examples of pseudo-Anosov maps and is not expected to be sharp. The examples we use are a subset of those considered by Farb, Leininger and Margalit, but we obtain better bounds for  $\ell_C(f)$ . The proof of the lower bound follows the same approach as Masur and Minsky [5], but by keeping track of more information, we obtain sharper bounds. The additional information comes from the algorithm of Bestvina and Handel that constructs an invariant train track for a pseudo-Anosov map.

The lower bound in Theorem 1.1 is a part of Theorem 5.1 in which we also prove a lower bound for punctured surfaces. To be precise, when  $\xi(S) \geq 2$  and  $n > 0$  we show that

$$\frac{1}{18(2g - 2 + n)^2 + 30(2g - 2 + n) - 10n} < L_C(\text{Mod}(S)).$$

At the end of the paper, we also discuss upper bounds for some families of punctured surfaces.

**Acknowledgements** We thank Chris Leininger for suggesting the project, and numerous discussions during the course of it. We thank the referee and the editor for helpful comments.

## 2 The complex of curves

For a broad range of questions in Teichmüller theory and mapping class groups, the coarse geometry of  $\mathcal{C}(S)$  plays a key role. See Minsky [6]. The curve complex  $\mathcal{C}(S)$

is quasi-isometric to its 1–skeleton equipped with the path metric. The 1–skeleton is a locally infinite graph. In [5, Proposition 4.6], Masur and Minsky showed that:

**Proposition 2.1** *For a nonsporadic surface  $S$ , there exists  $c > 0$  such that, for any pseudo-Anosov mapping class  $f$  and any simple closed curve  $\alpha$  in  $\mathcal{C}(S)$*

$$d_{\mathcal{C}}(f^n(\alpha), \alpha) \geq c|n|$$

for all  $n \in \mathbb{Z}$ .

In particular, the proposition shows that the curve complex  $\mathcal{C}(S)$  has infinite diameter. In the same paper, Masur and Minsky went on to show that the curve complex is  $\delta$ –hyperbolic in the sense of Gromov. Proposition 2.1 implies that pseudo-Anosov mapping classes have “north-south” dynamics on  $\mathcal{C}(S)$ , ie, they act as *hyperbolic* elements on  $\mathcal{C}(S)$  and have an invariant quasixis.

A consequence of Proposition 2.1 is that for nonsporadic surfaces  $S$ , the minimal asymptotic translation length  $L_{\mathcal{C}}(\text{Mod}(S)) > 0$ . In fact, Bowditch showed that the numbers  $\ell_{\mathcal{C}}(f)$  are rational with uniformly bounded denominators [2].

The following fact about the asymptotic lengths of iterates of  $f$  is useful for proving bounds.

**Lemma 2.2** *For all integers  $m \geq 1$ ,  $\ell_{\mathcal{C}}(f^m) = m\ell_{\mathcal{C}}(f)$ .*

**Proof** From the definition of  $\liminf$ ,

$$\begin{aligned} \ell_{\mathcal{C}}(f) &= \liminf_{j \rightarrow \infty} \frac{d_{\mathcal{C}}(\alpha, f^j(\alpha))}{j} \leq \liminf_{j \rightarrow \infty} \frac{d_{\mathcal{C}}(\alpha, f^{jm}(\alpha))}{jm} \\ &= \frac{1}{m} \liminf_{j \rightarrow \infty} \frac{d_{\mathcal{C}}(\alpha, f^{jm}(\alpha))}{j} = \frac{1}{m} \ell_{\mathcal{C}}(f^m). \end{aligned}$$

To get the reverse inequality, we use the triangle inequality, followed by the fact that  $f$  is an isometry of  $\mathcal{C}(S)$ , ie,

$$d_{\mathcal{C}}(\alpha, f^{jm}(\alpha)) \leq \sum_{i=1}^m d_{\mathcal{C}}(f^{(i-1)j}(\alpha), f^{ij}(\alpha)) = md_{\mathcal{C}}(\alpha, f^j(\alpha)).$$

Hence  $\ell_{\mathcal{C}}(f^m) \leq m\ell_{\mathcal{C}}(f)$  and we are done. □

### 3 Train tracks

For a detailed discussion of train tracks, see Penner and Harer [9]. We summarize the necessary definitions here.

A *train track*  $\tau$  on the surface is an embedded 1–dimensional CW complex with some additional structure. The edges are called *branches* and the vertices are called *switches*. The branches are smoothly embedded on the interiors, and there is a common point of tangency to all branches meeting at a switch. This splits the set of branches incident on a switch into two disjoint nonempty subsets, which can be arbitrarily assigned as the incoming and outgoing branches at the switch. In [5], Masur and Minsky assume all switches to have valence at least three. The particular results from [5] that we use here hold more generally with valence at least two.

A *train route* is a regular smooth path in  $\tau$ . In particular, it traverses a switch only by passing from an incoming edge to an outgoing edge or vice versa. A train track  $\sigma$  is *carried* by  $\tau$ , denoted by  $\sigma \prec \tau$ , if there is a homotopy of the identity map of the surface such that every train route in  $\sigma$  is taken to a train route in  $\tau$ . In particular, this means that  $\sigma$  can be embedded in an  $\epsilon$  neighborhood of  $\tau$ . A simple closed curve is *carried* by a train track if it is homotopic to a closed train route.

An assignment of nonnegative numbers, called *weights*, to the branches so that at every switch, the sum of the incoming weights equals the sum of the outgoing weights is called a *transverse measure* on the train track. A closed train route induces a counting measure on  $\tau$ .

Following Masur and Minsky [5], we shall denote the set  $P(\tau)$  to be the polyhedron of transverse measures supported on  $\tau$  and let  $\text{int}(P(\tau)) \subset P(\tau)$  be the set of transverse measures on  $\tau$  which induce positive weights on every branch of  $\tau$ . A simple closed curve  $\alpha$  carried by  $\tau$  naturally induces a transverse measure supported on  $\tau$ , so by abuse of notation  $\alpha \in P(\tau)$ .

A train track is called *large* if all the complementary regions are polygons or once-punctured polygons. A train track that has complementary regions ideal triangles or once-punctured monogons is called *maximal* or *complete*. It is maximal in the sense that it cannot be a subtrack of some other train track. The complementary regions of a large train track are *ideal* in the sense that the internal angle at all their vertices is zero. Hence, the vertices will be called *cusps*.

A train track  $\tau$  is called *recurrent* if there is a transverse measure which is positive on every branch of  $\tau$ . A train track  $\tau$  is *transversely recurrent* if given a branch of  $\tau$  there is a simple closed curve on  $S$  that crosses the branch and intersects  $\tau$  transversely and

efficiently, ie the union of  $\tau$  and the simple closed curve has no complementary bigons. A train track that is both recurrent and transversely recurrent is *birecurrent*.

A train track  $\sigma$  *fills*  $\tau$  if it is carried by  $\tau$  and  $\text{int}(P(\sigma)) \subset \text{int}(P(\tau))$ . For recurrent train tracks, this means that every branch of  $\tau$  is traversed by some branch of  $\sigma$ .

For a large train track  $\tau$ , a train track  $\sigma$  is called a *diagonal extension* of  $\tau$  if  $\tau$  is subtrack of  $\sigma$ , and each branch in  $\sigma \setminus \tau$  has its endpoints terminate in the cusps of a complementary regions of  $\tau$ . Let  $E(\tau)$  denote the set of all recurrent diagonal extensions of  $\tau$ . It is obvious that this is a finite set. Following Masur and Minsky, set

$$PE(\tau) = \bigcup_{\sigma \in E(\tau)} P(\sigma)$$

Further, let  $\text{int}(PE(\tau))$  be the set of measures in  $PE(\tau)$  that are positive on every branch of  $\tau$ .

We begin with a preliminary lemma of Masur and Minsky which will be useful in Section 5; the proof of the lemma can be found in [5].

**Lemma 3.1** [5] *For large recurrent train tracks  $\sigma, \tau$ , if  $\sigma$  fills  $\tau$ , then any  $\sigma' \in E(\sigma)$  is carried by some  $\tau' \in E(\tau)$ . In particular, there is the inclusion  $PE(\sigma) \subseteq PE(\tau)$ .*

### The nesting lemma of Masur and Minsky

Given a set  $A$  in  $\mathcal{C}(S)$ , let  $\mathcal{N}_1(A)$  denote the 1-neighborhood of  $A$  in  $\mathcal{C}(S)$ . In [5], Masur and Minsky showed the following important result:

**Lemma 3.2** (Nesting lemma) *Let  $\tau$  be a large birecurrent train track. Then*

$$\mathcal{N}_1(\text{int}(PE(\tau))) \subset PE(\tau).$$

In other words, if  $\alpha$  is a curve carried by a diagonal extension of  $\tau$  such that  $\alpha$  passes through every branch of  $\tau$ , and  $\beta$  is a curve disjoint from  $\alpha$  then  $\beta$  is also carried by some diagonal extension of  $\tau$ .

The original lemma in Masur and Minsky requires that  $\tau$  be birecurrent. We show below that the hypothesis of transverse recurrence can be dropped. The proof here was suggested by Chris Leininger.

**Proof** Let  $\tau$  be a large recurrent train track and let  $\alpha$  be a curve in  $\text{int}(PE(\tau))$ . Let  $\sigma$  be a diagonal extension of  $\tau$  carrying  $\alpha$  such that  $\alpha$  passes over every branch of  $\sigma$ , ie, to get  $\sigma$  we add to  $\tau$  only as many diagonals as necessary. Thus,  $\alpha \in \text{int}(P(\sigma))$ . We claim:

**Claim 1** *Let  $\beta$  be a curve disjoint from  $\alpha$ . Then,  $\beta \in PE(\sigma)$ .*

**Proof** For each branch  $b$  of  $\sigma$ , denote by  $\alpha(b)$  the weight assigned to  $b$  by  $\alpha$ . Since  $\alpha \in \text{int}(P(\sigma))$ , the weights  $\alpha(b) > 0$  for all  $b$ .

Consider  $\sigma$  as an abstract train track. To each branch  $b$ , assign a rectangle  $R(b)$ , of length 1 and width  $\alpha(b)$ . Foliate each rectangle by the product foliations, ie, by horizontal and vertical lines. The weights  $\alpha(b)$  satisfy the switch conditions. So the rectangles glue along their widths in a consistent manner to give a neighborhood  $\mathcal{N}$  of  $\sigma$ . See Figure 1.

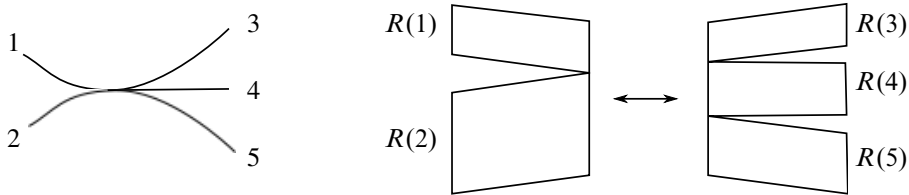


Figure 1: Gluing rectangles

The foliations also glue up to give a pair of singular foliations of  $\mathcal{N}$ , which we continue to call horizontal and vertical. The horizontal foliation is obtained from a cylinder neighborhood of  $\alpha$ , foliated by leaves parallel to  $\alpha$ , and with parts of its boundary glued together. In particular, we view  $\alpha$  as a leaf in the horizontal foliation of  $\mathcal{N}$ . The vertical foliations is by *ties* for  $\sigma$ . The components of the boundary  $\partial\mathcal{N}$ , are each a finite union of arcs of singular leaves of the horizontal foliation, and correspond precisely to the complementary polygons of  $\sigma$ . In fact,  $\mathcal{N}$  admits an embedding into the surface  $S$  as a neighborhood of  $\sigma$ .

The union of vertical sides of all rectangles is a union of leaves of the vertical foliation. Denote this union as  $L$ . Let  $Z_1, \dots, Z_u$  be the complementary polygons of  $\sigma$ . The key observation is that each  $Z_i$  that has, say,  $k$  sides is contained in a unique  $2k$ -gon  $Y_i$  whose sides in a cyclic order, are alternatively arcs of  $L$  and arcs of  $\alpha$ . For instance, in Figure 2, the complementary region  $Z_i$  is a quadrilateral. As shown in the picture on the left, a side of  $Z_i$  is a union of branches. The rectangles corresponding to the branches are shown in the picture on the right. The dotted lines in the union of these rectangles are the sides of the  $2k$ -gon that are arcs of  $\alpha$ . Each of these arcs is a piece of  $\alpha$  first encountered as we move out from  $Z_i$  across its sides.

The surface  $S$  decomposes into a union of the even-gons and a set of rectangles  $X_1, \dots, X_v$ . For each  $X_i$ , a pair of opposite sides are arcs of  $L$ , and the other pair of

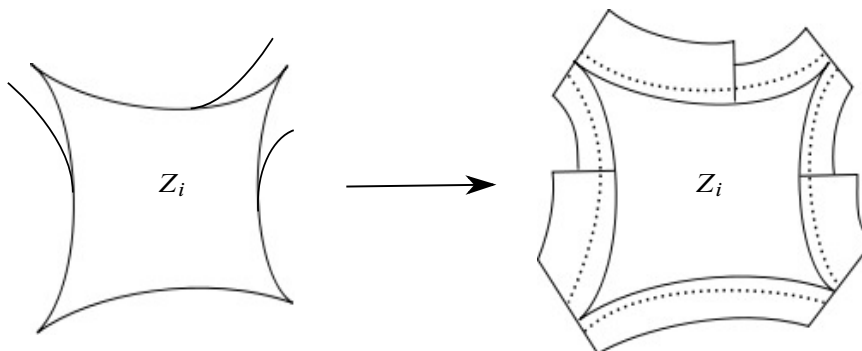


Figure 2: The  $2k$ -gon

opposite sides are arcs of  $\alpha$ . Also, each  $X_i$  is contained in an original rectangle  $R(b_i)$ , and that  $R(b_i)$  contains no other  $X_k$ .

Keeping  $\beta$  disjoint from  $\alpha$ , we isotope  $\beta$  to minimize the number of intersection points with  $L$ . By construction, each arc in  $\beta \setminus \beta \cap L$  has to be contained entirely in either a single rectangle  $X_i$  or a single even-gon  $Y_j$ , and must connect a  $L$ -side to another  $L$ -side. An arc inside  $X_i$  connecting its  $L$ -sides traverses the branch  $b_i$ . If an arc in  $Y_j$  connects consecutive  $L$ -sides in a cyclic order on the  $L$ -sides of  $Y_j$ , then it traverses a side of  $Z_j$ , which is a union of branches of  $\sigma$ . On the other hand, if an arc in  $Y_j$  connects nonconsecutive  $L$ -sides, then it traverses a diagonal of  $Z_j$ .

It follows that  $\beta$  is carried by a diagonal extension of  $\sigma$ , proving the claim. □

Diagonal extensions of  $\sigma$  are also diagonal extensions of  $\tau$ . So, the claim implies that  $\beta \in PE(\tau)$ , finishing the proof of Lemma 3.2. □

## 4 The Bestvina–Handel algorithm

**Definition 4.1** Given a pseudo-Anosov mapping class  $f \in \text{Mod}(S)$ , a train track  $\tau$  is an invariant train track of  $f$  if  $\tau$  is large and recurrent, and  $f(\tau) < \tau$ .

**The Bestvina–Handel train track** The Bestvina–Handel algorithm takes as input a punctured surface with a pseudo-Anosov map  $f$  and constructs an invariant train track  $\tau$  for  $f$ . The algorithm extends to closed surfaces as follows: Given a pseudo-Anosov map  $f$  of a closed surface, a singularity of the stable foliation has a finite orbit under  $f$ . After puncturing the surface at these orbit points, the map  $f$  restricts to a

pseudo-Anosov map of the punctured surface. Running the Bestvina–Handel algorithm for the punctured surface yields a train track  $\tau$  that is also an invariant train track for the closed surface. For details about the algorithm, we refer to Bestvina and Handel’s paper [1]. Here, we present the features of the track  $\tau$  that we need in the proof of Theorem 5.1.

(1) The branches of  $\tau$  are essentially of two types: *real* and *infinitesimal*. The reason for this classification is that in passing from  $\tau$  to the associated Markov partition for the stable foliation, only the real branches correspond to rectangles. For a surface with punctures, the algorithm can be carried out such that the number  $r$  of real branches of  $\tau$  is bounded above by  $3|\chi(S)|$ , where  $\chi(S)$  is the Euler characteristic of  $S$ . For a closed surface, the number of singularities of a measured foliation is bounded above by  $2|\chi(S)|$ . This implies  $r \leq 9|\chi(S)|$  for the number of real branches on a closed surface  $S$ .

We also need an upper bound for the number of infinitesimal branches. By [1, Propositions 3.3.3 and 3.3.4], the number of valence–2 switches in a Bestvina–Handel train track is at most twice the number of cusps. Furthermore, each valence–2 switch occurs either between an infinitesimal branch and a real branch or between two infinitesimal branches. The number of cusps is the largest when a train track is maximal, and is equal to  $6|\chi(S)| - 2n$ . This implies the crude upper bound  $24|\chi(S)| - 8n$  for the number of infinitesimal branches.

(2) Along with the track  $\tau$ , the algorithm gives a map  $\tau \rightarrow \tau$  taking switches to switches, that is efficient in a certain sense. See [1, Lemma 3.1.2]. There is a homotopy of  $S$  sending  $f(\tau)$  into  $\tau$  such that the resulting map  $\tau \rightarrow \tau$  is the one above. By abuse of notation, we will denote the map from  $\tau$  to itself also by  $f$ . Since  $f$  maps switches to switches, there is an unambiguously defined transition matrix  $M$  with entries corresponding to ordered pairs of branches in  $\tau$  such that the entry corresponding to the pair  $(b_1, b_2)$  counts the number of times  $f(b_2)$  passes over  $b_1$ . Also because  $f$  maps switches to switches, the transition matrix for  $f^k$  is  $M^k$ .

**Definition 4.2** A matrix  $M$  is *nonnegative* if every entry in it is nonnegative. A nonnegative matrix  $M$  is

- (a) *irreducible* if for any  $(i, j)$ , there exist a positive integer  $s$  such that  $M^s$  has a positive  $(i, j)$ -th entry.
- (b) *primitive* if it is irreducible and nonnegative, and  $M^s$  is a positive matrix for some  $s$ .

Bestvina and Handel show that the  $r \times r$  submatrix  $M_{\mathcal{R}}$  of  $M$  obtained by restricting to the set of real branches  $\mathcal{R}$  is irreducible. In fact, as will be clear in the proof of



Lemma 4.3,  $M_{\mathcal{R}}$  is primitive. Additionally, for every infinitesimal branch in  $\tau$ , there is a real branch such that some iterate of it passes over the infinitesimal branch.

With  $n$  denoting the number of punctures, set  $c_0 = 162$  and  $c_n = 18$  for all  $n > 0$ . A consequence of the discussion above is:

**Lemma 4.3** *There is positive integer  $k < c_n \chi(S)^2 + 24|\chi(S)| - 8n$  such that for any real branch  $b$ , the path  $f^k(b)$  passes over every branch of  $\tau$ .*

**Proof** In [1], Bestvina and Handel show that the transition matrix  $M$  has the form

$$(4-1) \quad M = \begin{pmatrix} A & B \\ 0 & M_{\mathcal{R}} \end{pmatrix},$$

where  $A$  is a permutation matrix. In other words,  $f$  permutes the infinitesimal branches.

By the definition of irreducibility, some iterate  $(M_{\mathcal{R}})^q$  has a positive diagonal entry. In fact, irreducibility implies path connectivity of the (directed) adjacency graph  $G$  associated to  $M_{\mathcal{R}}$ . Any directed path of length  $(r + 1)$  in  $G$  must necessarily contain a loop. Since  $\text{Tr}((M_{\mathcal{R}})^q)$  counts the number of loops of length  $q$  in  $G$ , it follows that  $q \leq r$ . See Gantmacher [4] or Seneta [10]. From (4-1), it follows that the  $r \times r$  submatrix  $M_{\mathcal{R}}^q$  of  $M^q$  given by the restriction to  $\mathcal{R}$  is  $(M_{\mathcal{R}})^q$ , ie,  $M_{\mathcal{R}}^q = (M_{\mathcal{R}})^q$ . The matrix  $M^q$  is the transition matrix for the iterate  $f^q$  which is pseudo-Anosov. So  $M_{\mathcal{R}}^q$  is still irreducible. By [11, Proposition 2.4] applied to  $M_{\mathcal{R}}^q$ , we know that  $(M_{\mathcal{R}}^q)^{2r}$  is a positive matrix. In particular,  $M_{\mathcal{R}}$  is primitive.

If we set  $p = 2rq$ , then for any real branch  $b \in \mathcal{R}$ , the path  $f^p(b)$  passes over all real branches. This implies that  $\mathcal{R} \subseteq f(\mathcal{R})$ . Since every infinitesimal branch is passed over by an iterate of some real branch, the inclusion is strict. By iterations, we get the sequence of inclusions  $\mathcal{R} \subset f(\mathcal{R}) \subset f^2(\mathcal{R}) \subset \dots$ , where the inclusion remains strict as long as  $f^{j+1}(\mathcal{R})$  spreads over a larger set of infinitesimal branches than  $f^j(\mathcal{R})$ . Let  $i$  be the smallest positive integer such that the sequence stabilizes, ie,  $f^i(\mathcal{R}) = f^{i+1}(\mathcal{R})$ . Then  $f^i(\mathcal{R}) = \tau$  and the number  $i$  is bounded above by the number of infinitesimal branches, which in turn is bounded above by  $24|\chi(S)| - 8n$ .

Set  $k = p + i$ . Then, for any real branch  $b \in \mathcal{R}$ , the path  $f^k(b)$  crosses all branches of  $\tau$ . It remains to give an upper bound for  $k$  in terms of  $\chi(S)$ .

For a nonsporadic surface with punctures,  $0 < r \leq 3|\chi(S)|$ . So we get the bound

$$k = p + i \leq 2r^2 + i \leq 18\chi(S)^2 + 24|\chi(S)| - 8n.$$

For a nonsporadic closed surface,  $0 < r \leq 9|\chi(S)|$ . So we get the bound

$$k = p + i \leq 2r^2 + i \leq 162\chi(S)^2 + 24|\chi(S)|. \quad \square$$

(3) We shall regard a cusp of a complementary region of  $\tau$  as *foldable*, if the branches  $b_1$  and  $b_2$  that flank it *fold* under some iterate, ie, if we orient  $b_1$  and  $b_2$  away from the cusp, there is some iterate such that the paths  $f^j(b_1)$  and  $f^j(b_2)$  starting from the same switch pass over the same initial branch  $b$ . By [1, Property (I2), Section 4], the algorithm is carried out such that the branch  $b$  that they fold over is always real. Let  $\sigma \in E(\tau)$ , and let  $\gamma$  be a simple closed curve carried by  $\sigma$ . We have the following lemma:

**Lemma 4.4** *If  $\gamma$  does not pass over any real branch of  $\tau$ , then  $\gamma$  is incident on a foldable cusp.*

**Proof** Suppose that  $\gamma$  does not pass over any real branch of  $\tau$  and none of the cusps it passes through are foldable. Any iterate of  $\gamma$  must also have the same properties. But then, the iterates cannot converge to the stable foliation of  $f$ , giving a contradiction.  $\square$

## 5 Lower bounds

**Theorem 5.1** *When the complexity  $\xi(S) \geq 2$ ,*

$$L_C(\text{Mod}(S)) > \frac{1}{c_n \chi(S)^2 + 30|\chi(S)| - 10n},$$

where, as in Lemma 4.3, the constants are  $c_0 = 162$  and  $c_n = 18$  for  $n \geq 1$ .

**Proof** The idea is to combine Lemma 4.3 with the proof [5, Proposition 4.6] to obtain a better lower bound.

For any pseudo-Anosov  $f \in \text{Mod}(S)$ , let  $\tau$  be the invariant train track of  $f$  constructed by Bestvina and Handel. We shall show that after at most  $(6|\chi(S)| - 2n + k)$  iterates of  $f$ , where  $k$  is the number of iterates in Lemma 4.3, we get the nesting behavior in the proof of Proposition 4.6 in [5]. This gives the lower bound for  $L_C(\text{Mod}(S))$  as stated.

For the track  $\tau$ , let  $\mathcal{B}_\tau$  be the set of the branches, and  $|\mathcal{B}_\tau|$  its cardinality. Let  $\sigma \in E(\tau)$ . By Lemma 3.1,  $PE(f(\tau)) \subset PE(\tau)$ . So the image  $f(\sigma)$  is carried by some diagonal extension  $\sigma' \in E(\tau)$ . Moreover,  $f$  sends switches of  $\sigma$  to switches of  $\sigma'$ . Hence, for each such pair  $(\sigma, \sigma')$ , the transition matrix  $M_{\sigma, \sigma'}: \mathbb{R}^{\mathcal{B}_\sigma} \rightarrow \mathbb{R}^{\mathcal{B}_{\sigma'}}$  associated to  $f$  is unambiguously defined. Without loss generality, we may assume that the last  $|\mathcal{B}_\tau|$  coordinates correspond to  $\mathcal{B}_\tau$ . Then, the matrix  $M_{\sigma, \sigma'}$  has the form

$$M_{\sigma, \sigma'} = \left( \begin{array}{c|c} * & 0 \\ \hline * & M \end{array} \right),$$

where  $M$  is the transition matrix for  $\tau$  described in Lemma 4.3. For any diagonal extension  $\sigma_0 \in E(\tau)$ , for any  $m > 0$ , we can construct a sequence of train tracks  $\sigma_1, \sigma_2, \dots, \sigma_m$  in  $E(\tau)$  such that

$$f(\sigma_0) \prec \sigma_1, f(\sigma_1) \prec \sigma_2, \dots, f(\sigma_{m-1}) \prec \sigma_m,$$

hence

$$f^m(\sigma_0) \prec \sigma_m.$$

Let  $M_{\sigma_j, \sigma_{j+1}}: \mathbb{R}^{\mathcal{B}_{\sigma_j}} \rightarrow \mathbb{R}^{\mathcal{B}_{\sigma_{j+1}}}$  be the transition matrices of  $f$  associated to  $f(\sigma_j) \prec \sigma_{j+1}$  in the sequence, and let  $M_{\sigma_0, \sigma_m}: \mathbb{R}^{\mathcal{B}_{\sigma_0}} \rightarrow \mathbb{R}^{\mathcal{B}_{\sigma_m}}$  be the transition matrix associated to  $f^m$ . Because  $f$  maps switches to switches, the matrices satisfy

$$\begin{aligned} M_{\sigma_0, \sigma_m} &= M_{\sigma_{m-1}, \sigma_m} \times M_{\sigma_{m-2}, \sigma_{m-1}} \times \dots \times M_{\sigma_1, \sigma_2} \times M_{\sigma_0, \sigma_1} \\ &= \left( \begin{array}{c|c} * & 0 \\ \hline * & M \end{array} \right) \times \left( \begin{array}{c|c} * & 0 \\ \hline * & M \end{array} \right) \times \dots \times \left( \begin{array}{c|c} * & 0 \\ \hline * & M \end{array} \right) \times \left( \begin{array}{c|c} * & 0 \\ \hline * & M \end{array} \right) = \left( \begin{array}{c|c} * & 0 \\ \hline * & M^m \end{array} \right). \end{aligned}$$

Let  $\mu \in P(\sigma_0)$ . We denote by  $f^m(\mu)$  the measure in  $P(\sigma_m)$  given by the pushforward of  $\mu$  by  $M_{\sigma_0, \sigma_m}$ . We now use Lemma 4.3 and Lemma 4.4 to prove the following lemma:

**Lemma 5.2** *For any  $\mu \in P(\sigma_0)$ , there exists some positive integer  $m$  such that  $k \leq m \leq 6|\chi(S)| - 2n + k$ , where  $k$  is the number of iterates in Lemma 4.3, the measure  $f^m(\mu) \in P(\sigma_m)$  is positive on every branch in  $\mathcal{B}_\tau$ , that is  $f^m(\mu) \in \text{int}(PE(\tau))$ .*

**Proof** We consider the simplest case first:

**Case 1** Suppose  $\mu$  is positive on some real branch  $b$  in  $\mathcal{B}_\tau$ . By Lemma 4.3, the transition matrix with respect to  $\tau$ , for  $f^k$  has the form

$$M^k = \begin{pmatrix} * & \widehat{B} \\ 0 & M_{\mathcal{R}}^k \end{pmatrix},$$

where  $\widehat{B}$  and  $M_{\mathcal{R}}^k$  are positive matrices. In particular, the image path  $f^k(b)$  passes over every branch in  $\mathcal{B}_\tau$ . Hence, the measure  $f^k(\mu) = M_{\sigma_0, \sigma_k}(\mu)$  in  $P(\sigma_k)$  is positive on every branch in  $\mathcal{B}_\tau$ . The same reasoning applied to all integers  $m \geq k$  implies that the measure  $f^m(\mu) = M_{\sigma_0, \sigma_m}(\mu)$  is positive on every branch of  $\mathcal{B}_\tau$ , finishing the proof of Lemma 5.2 in this case.

**Case 2** Suppose  $\mu$  is not positive on any real branch. We shall show that in  $j \leq 6|\chi(S)| - 2n$  iterates the measure  $f^j(\mu) = M_{\sigma_0, \sigma_j}(\mu)$  is positive on some real branch, reducing us to Case 1. This is done in two steps: In Step 1, we show that  $\text{Supp}(\mu)$  contains a diagonal  $d$  that is incident on a foldable cusp  $c$ . In Step 2, we show that

the branches  $b_1$  and  $b_2$  that flank  $c$ , fold over a real branch  $b$  in  $j \leq 6|\chi(S)| - 2n$  iterates. Then  $f^j(d)$  also passes over  $b$  from which it follows that  $f^j(\mu)$  assigns positive weight to  $b$ .

**Step 1** Suppose  $\mu$  is positive on some simple closed curve  $\gamma$  carried by  $\sigma_0$ . By Lemma 4.4, the curve  $\gamma$  must be incident on a foldable cusp. Hence,  $\text{Supp}(\mu)$  contains a diagonal  $d$  that is incident on a foldable cusp  $c$ .

**Step 2** Let  $b_1$  and  $b_2$  be the branches that flank  $c$ . Let  $j$  be the smallest iterate in which  $b_1$  and  $b_2$  fold. By Part (3) of Section 4, the branch  $b$  that they fold over is real. We claim that  $j \leq 6|\chi(S)| - 2n$ . By an Euler characteristic calculation, the total number of cusps is at most  $6|\chi(S)| - 2n$ . If  $b_1$  and  $b_2$  do not fold within  $6|\chi(S)| - 2n$  iterates, then there is a foldable cusp  $c'$  such that  $f^a(c') = c'$  for some iterate  $f^a$ . But then  $f^{ma}(c') = c'$  for all positive integers  $m$ . Thus,  $c'$  never gets folded giving a contradiction. This proves the claim.

Combining this with Case 1, we conclude that for  $m = j + k \leq 6|\chi(S)| - 2n + k$ , the measure  $f^m(\mu)$  is positive on every branch in  $\mathcal{B}_\tau$  finishing the proof of Lemma 5.2.  $\square$

Back to Theorem 5.1, Lemma 5.2 implies that for any  $\sigma_0 \in E(\tau)$ , and for any  $\mu \in P(\sigma_0)$ ,

$$f^w(\mu) \in \text{int}(PE(\tau)),$$

where  $w = 6|\chi(S)| - 2n + k$ . Hence,

$$(5-1) \quad f^w(PE(\tau)) \subset \text{int}(PE(\tau)).$$

Now set  $\tau_1 = \tau$ , and for each positive integer  $i \geq 1$ , let  $\tau_i = f^{iw}(\tau)$ . The inclusion (5-1) implies  $PE(\tau_{i+1}) \subset \text{int}(PE(\tau_i))$ . By Lemma 3.2, we get the nesting sequence:

$$P(\tau_{i+1}) \subset \text{int}(PE(\tau_i)) \subset \mathcal{N}_1(\text{int}(PE(\tau_i))) \subset PE(\tau_i) \subset \dots \\ \subset \text{int}(PE(\tau_1)) \subset \mathcal{N}_1(\text{int}(PE(\tau_1))) \subset PE(\tau_1)$$

Choose  $\alpha \in \mathcal{C}(S) \setminus PE(\tau_1)$  such that  $f^w(\alpha) \in PE(\tau_1)$ . Then  $f^{iw}(\alpha)$  is in  $PE(\tau_i)$  but not in  $PE(\tau_{i+1})$ . Thus  $d_{\mathcal{C}}(f^{iw}(\alpha), \alpha) \geq i$ . Hence

$$\ell_{\mathcal{C}}(f^w) = \liminf_{i \rightarrow \infty} \frac{d_{\mathcal{C}}(f^{iw}(\alpha), \alpha)}{i} \geq \liminf_{i \rightarrow \infty} \frac{i}{i} = 1.$$

By Lemma 2.2, we have  $\ell_{\mathcal{C}}(f^w) = w\ell_{\mathcal{C}}(f)$ . So

$$\ell_{\mathcal{C}}(f) \geq \frac{1}{w} > \frac{1}{c_n\chi(S)^2 + 30|\chi(S)| - 10n},$$

where  $c_0 = 162$  and  $c_n = 18$  for  $n \geq 1$ .  $\square$

## 6 Upper bound

Next, for a closed surface  $S$ , we prove an upper bound for  $L_C(\text{Mod}(S))$  of the same order.

**Theorem 6.1** For a closed surface of genus  $g \geq 2$ ,

$$L_C(\text{Mod}(S)) \leq \frac{4}{g^2 + g - 4}.$$

**Proof** It is sufficient to find a pseudo-Anosov mapping class  $f$  such that  $\ell_C(f) \leq 4/(g^2 + g - 2)$ . We show this for the pseudo-Anosov map of a closed surface of genus  $g$  constructed by Penner in [8]. The Penner example is as follows: For the closed surface of genus  $g$  in Figure 3, let  $f = \rho T_{c_1} T_{b_1}^{-1} T_{a_1}$ , where  $T_{a_1}$  is a positive Dehn twist along  $a_1$ ,  $\rho(a_i) = a_{i-1}$ , for  $i = 2, \dots, g$  and  $\rho(a_1) = a_g$  and similarly for the  $b_i$ 's and  $c_i$ 's.

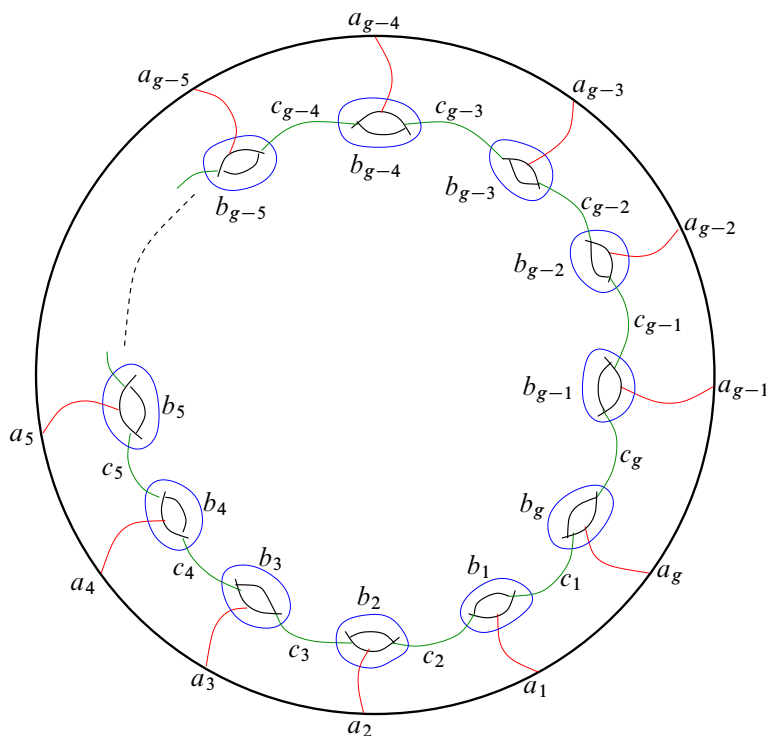


Figure 3:  $f = \rho T_{c_1} T_{b_1}^{-1} T_{a_1} \in \text{Mod}(S)$

Since  $\ell_C(f)$  is independent of the initial choice of curve to apply iterations to, we choose the curve  $a_g$  and show that for some  $k \geq (g^2 + g - 4)/2$ ,

$$(6-1) \quad d_C(f^k(a_g), a_g) \leq 2.$$

By the triangle inequality,

$$\ell_C(f^k) = \liminf_{j \rightarrow \infty} \frac{d_C(f^{jk}(a_g), a_g)}{j} \leq \liminf_{j \rightarrow \infty} \frac{2j}{j} = 2,$$

and by Lemma 2.2,

$$\ell_C(f) \leq \frac{2}{k} \leq \frac{4}{g^2 + g - 4}.$$

For  $g = 2$ , let  $k = 1 \geq (2^2 + 2 - 4)/2 = 1$  and notice that  $f(a_2) = a_1$ . Since  $a_1$  and  $a_2$  are disjoint,  $d_C(f(a_2), a_2) = 1 < 2$ , and we are done.

For a sequence of curves  $\alpha_r \in \{a_i, b_i, c_i\}_{i=1}^g$  such that  $\alpha_1 \cup \dots \cup \alpha_k$  is connected, we denote the regular neighborhood of the union  $\alpha_1 \cup \dots \cup \alpha_k$  by  $\mathcal{N}(\alpha_1 \dots \alpha_k)$ . To show (6-1) in general, the key idea is as follows: Suppose that  $f^j$  is the smallest iterate in which  $f^j(a_g)$  is spread over  $k$  ‘‘holes’’. Then it takes waiting time  $(g + 1)$  for the images to sweep over  $(k + 2)$  holes. In other words,  $f^{j(g+1)}$  is the smallest iterate in which the image of  $a_g$  sweeps over  $(k + 2)$  holes. To be precise, among the neighborhoods defined above, we keep track of which is the ‘‘smallest’’ one containing the image of  $a_g$ .

In first  $(g - 1)$  iterates  $a_g$  gets rotated till it becomes  $a_1$ , ie,  $f^{g-1}(a_g) = a_1$ . In two iterates that follow,

$$f^g(a_g) \subset \mathcal{N}(a_g b_g c_g), \quad f^{g+1}(a_g) \subset \mathcal{N}(c_g a_{g-1} b_{g-1} c_{g-1}).$$

In the same manner, continuing the iterations, notice that

$$f^{2(g+1)}(a_g) \subset \mathcal{N}(c_g b_{g-1} c_{g-1} a_{g-2} b_{g-2} c_{g-2} b_{g-3} c_{g-3}).$$

$$f^{3(g+1)}(a_g) \subset \mathcal{N}(c_g b_{g-1} c_{g-1} a_{g-2} b_{g-2} c_{g-2} \dots a_{g-4} b_{g-4} c_{g-4} b_{g-5} c_{g-5}).$$

We observe that after each  $f^{g+1}$  iterates the subscript for  $c$  rightmost inside  $\mathcal{N}$  decreases by 2. In other words, it requires  $(g + 1)$  iterates to increase the ‘‘complexity’’ of the image of  $a_g$  by 2. Here, we abbreviate notation by setting  $\mathcal{N}(c_g * c_{g-2i+1})$  to be  $\mathcal{N}(c_g b_{g-1} c_{g-1} a_{g-2} b_{g-2} c_{g-2} \dots a_{g-2i+2} b_{g-2i+2} c_{g-2i+2} b_{g-2i+1} c_{g-2i+1} a_{g-i})$ .

Then,

$$\begin{aligned}
 f^{g+1}(a_g) &\subset \mathcal{N}(c_g * c_{g-1}), \\
 f^{2(g+1)}(a_g) &\subset \mathcal{N}(c_g * c_{g-3}) \\
 f^{3(g+1)}(a_g) &\subset \mathcal{N}(c_g * c_{g-5}) \\
 &\vdots \\
 f^{\lfloor (g-1)/2 \rfloor (g+1)}(a_g) &\subset \mathcal{N}(c_g * c_{g-2\lfloor (g-1)/2 \rfloor + 1}).
 \end{aligned}$$

Finally,  $f^{g-1}(f^{\lfloor (g-1)/2 \rfloor (g+1)}(a_g)) \subset \mathcal{N}(\rho^{-1}(c_g * c_{g-2\lfloor (g-1)/2 \rfloor + 1})b_3c_3b_2c_2)$ ,

where the inclusion is into a smaller neighborhood if  $g$  is even.

Notice  $\mathcal{N}(\rho^{-1}(c_g * c_{g-2\lfloor (g-1)/2 \rfloor + 1})b_3c_3b_2c_2)$ , and so  $f^{g-1}(f^{\lfloor (g-1)/2 \rfloor (g+1)}(a_g))$  is disjoint from  $a_1$ , and of course  $a_1$  and  $a_g$  are disjoint. Hence

$$d_C(a_g, f^k(a_g)) \leq 2,$$

where

$$k = (g - 1) + \left\lfloor \frac{g - 1}{2} \right\rfloor (g + 1) \geq \frac{2(g - 1) + (g - 2)(g + 1)}{2} = \frac{g^2 + g - 4}{2}. \quad \square$$

More generally, as described in [11, Appendix 5.2], a method similar to [8] constructs pseudo-Anosov homeomorphisms of certain punctured surfaces from pseudo-Anosov homeomorphisms of closed surfaces. We start with the Penner pseudo-Anosov map  $f$  of the closed surface  $S_g$ . We add in punctures in some or all of the complementary regions according to the criteria of [7, Theorem 3.1]. Then, the restriction of  $f$  is a pseudo-Anosov on the punctured surface. A proof similar to Theorem 6.1 provides upper bounds on  $L_C(\text{Mod}(S_{g,n}))$  of the order  $1/\chi(S_{g,n})^2$ . We list the cases in which we get  $1/\chi(S_{g,n})^2$  type upper bounds:

- (1) For punctured tori with  $n$  even, we use the example in [11, Appendix 5.1].
- (2) For  $g \geq 5$  and  $n = g - 1$  or  $2g - 2$ , we use [11, Example 3, Appendix 5.2].
- (3) For  $g \geq 3$  and  $n \leq 4$ , we use of [11, Example 2, Appendix 5.2].
- (4) For  $g \geq 2$  and  $n = 1, 2, g, g + 1$  or  $g + 2$ , we use Penner’s example in Theorem 6.1, puncturing the surface at the appropriate points.

In some cases, the upper bound can be of the order of  $1/|\chi(S_{g,n})|$ . For example, when  $g = 2$  and  $n$  is varying, the example of [11, Section 4] gives the bound

$$L_C(\text{Mod}(S_{g,n})) \leq \frac{20}{n - 4},$$

for all  $n \geq 4$ . We propose the following conjecture:

**Conjecture 6.2** For fixed  $g \geq 2$  and  $n$  varying,  $L_C(\text{Mod}(S_{g,n}))$  is of the order of  $1/|\chi(S_{g,n})|$  as  $n \rightarrow \infty$ .

## References

- [1] **M Bestvina, M Handel**, *Train-tracks for surface homeomorphisms*, *Topology* 34 (1995) 109–140 MR1308491
- [2] **B H Bowditch**, *Tight geodesics in the curve complex*, *Invent. Math.* 171 (2008) 281–300 MR2367021
- [3] **B Farb, C J Leininger, D Margalit**, *The lower central series and pseudo-Anosov dilatations*, *Amer. J. Math.* 130 (2008) 799–827 MR2418928
- [4] **F R Gantmacher**, *The theory of matrices. Vols. 1, 2*, Translated by K A Hirsch, Chelsea Publishing Co., New York (1959) MR0107649
- [5] **H A Masur, Y N Minsky**, *Geometry of the complex of curves. I: Hyperbolicity*, *Invent. Math.* 138 (1999) 103–149 MR1714338
- [6] **Y N Minsky**, *Curve complexes, surfaces and 3-manifolds*, from: “International Congress of Mathematicians. Vol. II”, (M Sanz-Solé, J Soria, J L Varona, J Verdera, editors), *Eur. Math. Soc.* (2006) 1001–1033 MR2275633
- [7] **R C Penner**, *A construction of pseudo-Anosov homeomorphisms*, *Trans. Amer. Math. Soc.* 310 (1988) 179–197 MR930079
- [8] **R C Penner**, *Bounds on least dilatations*, *Proc. Amer. Math. Soc.* 113 (1991) 443–450 MR1068128
- [9] **R C Penner, J L Harer**, *Combinatorics of train tracks*, *Annals of Math. Studies* 125, Princeton Univ. Press (1992) MR1144770
- [10] **E Seneta**, *Nonnegative matrices and Markov chains*, second edition, Springer Series in Statistics, Springer, New York (1981) MR719544
- [11] **C-Y Tsai**, *The asymptotic behavior of least pseudo-Anosov dilatations*, *Geom. Topol.* 13 (2009) 2253–2278 MR2507119

*Department of Mathematics, Harvard University  
Cambridge MA 02138, USA*

*Department of Mathematics, University of Illinois Urbana-Champaign  
Urbana IL 61801, USA*

vaibhav@math.harvard.edu, cyt1230@gmail.com

<http://www.math.harvard.edu/~vaibhav>, <http://www.math.uiuc.edu/~ctsai6>

Proposed: Joan Birman

Received: 18 January 2011

Seconded: David Gabai, Walter Neumann

Revised: 18 June 2011