Forum of Mathematics, Sigma (2026), **Vol. 14**:e114 1–44
doi:10.1017/fms.2026.10255

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Quadratic-time computations for pseudo-anosov mapping classes

Dan Margalit[1], Balázs Strenner[2], Samuel J. Taylor[3] and S. Öykü Yurttaş[4]

[1]Vanderbilt University, USA; E-mail: dan.margalit@vanderbilt.edu (Corresponding author).
[2]Georgia Institute of Technology, Georgia; E-mail: strennerb@gmail.com.
[3]Temple University, USA; E-mail: samuel.taylor@temple.edu.
[4]Boğaziçi University, Istanbul, Türkiye; E-mail: saadetoyku.yurttas@bogazici.edu.tr.

Received: 28 December 2024 UTC; **Revised:** 25 December 2025 UTC; **Accepted:** 25 May 2026 UTC

2020 Mathematical Subject Classification: *Primary* – 57K20

Abstract

We give a quadratic-time algorithm to compute the stretch factor and the invariant measured foliations for a pseudo-Anosov element of the mapping class group. As input, the algorithm accepts a word (in any given finite generating set for the mapping class group) representing a pseudo-Anosov mapping class, and the length of the word is our measure of complexity for the input. The output is a train track and an integer matrix where the stretch factor is the largest real eigenvalue and the unstable foliation is given by the corresponding eigenvector. This is the first algorithm to compute stretch factors and measured foliations that is known to terminate in subexponential time.

Contents

1 Introduction **2**
1.1 The main technical result . . . 3
1.2 The algorithm . . . 10
1.3 New tools and an outline of the proof . . . 12
2 An example **15**
2.1 The ledger . . . 15
2.2 Applying the algorithm . . . 17
3 Slope **19**
3.1 Preliminaries . . . 20
3.2 Slope disparity lemmas . . . 21
3.3 Slopes and stretch factors . . . 24
3.4 Application to stable translation lengths in the curve graph . . . 25
4 Approximating rectangles **26**
4.1 The existence of approximating rectangles . . . 26
4.2 Small-slope geodesics and approximating rectangles . . . 30
5 The forcing lemma **31**
5.1 Left- and right-turning paths and slopes of train tracks . . . 32
5.2 Obstructing and agreeable branches . . . 33
5.3 Proof of the forcing lemma . . . 34

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

6 The fitting lemma **35**
6.1 The backfolding lemma 36
6.2 Proof of the fitting lemma 39
7 Proof of Theorem 1.1 **40**
References **43**

1. Introduction

In his famous survey paper from 1982, "Three dimensional manifolds, Kleinian groups and hyperbolic geometry," William Thurston laid out a series of 24 problems that would shape the field of low-dimensional topology for the ensuing decades [31]. Problem 20 is:

> *Develop a computer program to calculate the canonical form for a general diffeomorphism of a surface, and to calculate the action of the group of diffeomorphisms on $P\mathcal{L}_0(S)$. Use this to implement an algorithm for the word problem and the conjugacy problem in the group of isotopy classes of diffeomorphisms of a surface.*

Here, $P\mathcal{L}_0(S)$ is the space of (compactly supported) projective measured laminations on a surface S. One reason for Thurston's interest in this question is revealed by his Problem 21: *Develop a computer program to calculate hyperbolic structures on 3-manifolds.* Because of the central place of the theory of surfaces in low-dimensional topology, Problem 20 has importance in its own right.

Significant progress has been made on Thurston's problem (see below), but it has not been completely solved; there are algorithms to compute the canonical form in general, but not computer programs that work for all compact surfaces. In this paper, we give, for any compact surface, an easily implementable and efficient algorithm for computing the canonical form in the case where the diffeomorphism is assumed to be of a certain type (namely, pseudo-Anosov).

To state our main result requires some setup. Let $S = S_{g,p}$ be the surface obtained from the connected sum of g tori by deleting p points; the complexity of S is $\xi(S) = 3g + p - 3$. The Nielsen–Thurston classification theorem says that the elements of the mapping class group $\mathrm{Mod}(S)$ fall into three categories: periodic, reducible, and pseudo-Anosov, and that the last category is exclusive from the other two. To a pseudo-Anosov element f there are two pieces of data associated: a pair of measured foliations on S and a real number $\lambda > 1$. These are called the stable and unstable foliations of f and the stretch factor. Any foliation in S can be represented by a train track τ in S along with a vector in the vector space $V(\tau)$ spanned by the branches of τ.

A *Nielsen–Thurston package* for $f \in \mathrm{Mod}(S)$ is a pair (τ, D), where τ is a train track in S and D is a rational matrix that represents a linear map of $V(\tau)$, that has the stretch factor of f as its largest real eigenvalue, and that has the unstable foliation of f as a corresponding positive eigenvector.

Main Theorem. *Fix a surface $S = S_{g,p}$ and let Γ be any finite generating set for $\mathrm{Mod}(S)$. There is an explicit algorithm whose input is any word w in Γ that represents a pseudo-Anosov element of $\mathrm{Mod}(S)$, whose output is a Nielsen–Thurston package (τ, D) for that element, and whose running time is quadratic in the length of w and polynomial in the complexity $\xi(S)$.*

This main theorem is stated more explicitly as Corollary 1.2 below. We consider this result to be a step toward solving Thurston's problem in general. Our algorithm is in the spirit of Thurston's problem: we directly compute (part of) the action of a mapping class on the space of measured foliations on a surface S, which is canonically isomorphic to $P\mathcal{L}_0(S)$.

There is no other algorithm for computing stretch factors and foliations whose running time is known to be faster than exponential. Our paper does not address the issue of computing the eigendata for the matrix D, since there are many well-known (and fast) algorithms for this.

We can imagine many possible applications for our algorithm, for instance in determining the smallest stretch factor in $\mathrm{Mod}(S_g)$ for small g. Equally important are the various tools we develop in the paper. For example, with our theory of slopes (Section 3) and our forcing lemma (Proposition 5.1) we give a

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

new criterion for a curve to be carried by a train track. We also expect that our work can be used toward a fast algorithm for the conjugacy problem in the mapping class group.

Computing model and running time. To say that an algorithm runs in quadratic time (as opposed to polynomial time) requires us to specify our model. We work in the bit model, that is, we count bit operations on a deterministic multitape Turing machine (and we refer to the number of bit operations as the run time). Arithmetic is charged at bit cost: adding two t-bit integers costs $O(t)$ bit operations, and multiplying a t-bit integer by an s-bit integer costs $O(ts)$ bit operations (schoolbook multiplication). No unit-cost arithmetic is assumed.

Let n be the length of the word w in the main theorem and let $\xi = \xi(S)$. The heart of our algorithm is to multiply together $O(\xi^2 n)$ square matrices. The size of each matrix is $O(\xi)$ and the entries have fixed bit size. This takes $O(n^2 \xi^d)$ bit operations for some $d \leq 8$. The algorithm also uses $O(\xi)$ applications of the Zassenhaus algorithm whose running time is $O(n^2 \xi^d)$ for some $d \leq 9$.

The idea of our algorithm. The group $\mathrm{Mod}(S)$ acts in a piecewise-linear fashion on $\mathrm{MF}(S)$, the space of measured foliations on S. Our main technical theorem, Theorem 1.1, says that there is a number $Q = Q(S) = O(\xi^2)$, a cell structure on $\mathrm{MF}(S)$, and a point c in $\mathrm{MF}(S)$ so that for any pseudo-Anosov $f \in \mathrm{Mod}(S)$ the point $f^Q(c)$ lies in a region of $\mathrm{MF}(S)$ that has two properties: (1) it contains the unstable foliation for f and (2) the action of f on the region is linear. From these two properties we may readily extract the desired Nielsen–Thurston package, as the unstable foliation is an eigenvector of the linear map from item (2). We can compute $f^Q(c)$ and the action of f at $f^Q(c)$ in quadratic time; as above, this essentially reduces to the problem of multiplying together Qn matrices, where n is the word length of f, and the entries of the matrices are of bounded size.

Prior work. Our algorithm has its origins in work of Thurston and Mosher (as exposited by Casson–Bleiler [10, p. 146]), who gave a method for explicitly computing the action of $\mathrm{Mod}(S)$ on $\mathrm{MF}(S)$. In his famous research announcement, Thurston suggested that under iteration by a pseudo-Anosov mapping class, points in $\mathrm{MF}(S)$ converge to the unstable foliation "rather quickly." Our Theorem 1.1 is an affirmation of this.

Bell–Schleimer proved a result that, in their words, shows Thurston's suggestion to be false [4]. Their theorem says that, within the linear region described above, the convergence rate to the ray of unstable foliations is arbitrarily slow. In other words, for the corresponding matrices, the absolute value of the ratio of the two eigenvalues of largest absolute value can be arbitrarily close to 1. There is no contradiction: our result says that the sequence $c, f(c), f^2(c), \ldots$ quickly arrives at the linear region containing the unstable foliation, and the Bell–Schleimer result says that, once there, the convergence of the sequence to the ray of unstable foliations can be slow.

Other works have taken advantage of the fact that the sequence $f^k(c)$ eventually – and quickly – lands in the "right" linear region for f; see the work of Menzel–Parker [21], Moussafir [23], Finn–Thiffeault [13], Hall (who wrote the computer program Dynn) [16], the fourth-named author with Hall [17], and Bell (who wrote the computer programs flipper and curver) [6]. Our Theorem 1.1 is the first result to quantify how quickly the sequence $f^k(c)$ lands in the right linear region.

There are also algorithms to compute stretch factors and foliations that do not use the action on $\mathrm{MF}(S)$ directly. Most notably, Bestvina–Handel gave an algorithm that uses train tracks and is based on the Stallings theory of folding [7]. The expectation is that this algorithm has exponential time complexity. Brinkmann [9] implemented the algorithm for once-punctured surfaces in the computer program Xtrain and Hall implemented it for punctured disks in the program Trains.

1.1. The main technical result

In order to state our main technical result, Theorem 1.1, we require some notation and setup. As above, the surfaces in this paper are all of the form $S = S_{g,p}$. Before we start in earnest, we give an abstract topological definition that will be used in the discussion of the space of measured foliations.

Integral cone complexes. In this paper, an integral cone complex is a sort of cell complex where a cell is a positive cone in some $\mathbb{R}^d$ defined as the set of solutions to a finite set of integer linear equations and

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

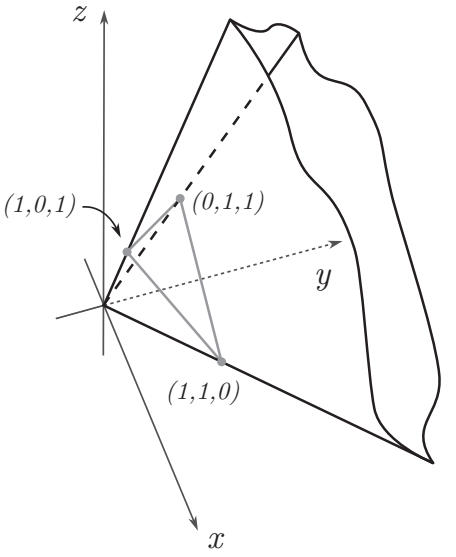


Figure 1. *A geometric realization of the integral cone complex Y_Δ.*

integer inequalities; we will refer to such a cone as an integral linear cone. Such a cell has a finite set of faces, each being a cell in the same sense. Cells are glued along their faces, the gluing maps being linear.

We now present a basic example of an integral cone complex Y_Δ that we will treat as a running example. Let V_1, V_2, and V_3 each be the first quadrant in $\mathbb{R}^2$. We glue the face of V_i on the x-axis to the face of V_{i+1} on the y-axis (subscripts considered modulo 3) by the map $(t, 0) \mapsto (0, t)$. The complex Y_Δ is homeomorphic to $\mathbb{R}^2$.

The integral cone complex Y_Δ can be geometrically realized in $\mathbb{R}^3$ as the boundary of the first octant (by a geometric realization, we mean an embedding into $\mathbb{R}^N$ mapping cells to cells). The complex Y_Δ is also geometrically realized as the set of points in the first octant of $\mathbb{R}^3$ whose coordinates satisfy a degenerate triangle inequality, meaning that one coordinate is the sum of the other two. See Figure 1 for an illustration.

With the definition of an integral cone complex in hand, we have several further definitions to make:

(a) integral linear maps, integral cone complex maps, and linear regions,
(b) the projectivized complex, vertex rays, and induced maps,
(c) the radius,
(d) subdivisions for generating sets,
(e) PL-eigenvectors,
(f) PL-eigenregions,
(g) projectively source-sink integral cone complex maps, and
(h) extended dynamical maps and sink packages.

After making these definitions, we explain how to compute PL-eigenvectors explicitly.

(a) In the category of integral cone complexes, the natural maps $f : Y \to Y$ are integral cone complex maps, which we now define. First, suppose U is a subset of some cell V of Y; we say that a map $U \to Y$ is integral linear if the image is contained in a cell W and – using the V- and W-coordinates – the map is given by the restriction of an integral matrix.

With this definition in hand, a map $f : Y \to Y$ is an integral cone complex map if there is an integral cone complex subdivision Y_f of Y so that the restriction of f to each cell of Y_f is an integral linear map. If V_i is a cell of Y we denote the cells of such a subdivision of V_i by $V^f_{i,k}$. We will refer to the cells of any such subdivision as linear pieces for f. These pieces are not canonical; the important feature for us is that each $V^f_{i,k}$ is determined by finitely many integer linear inequalities in the V_i-coordinates.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

If U is any integral linear cone contained in a cell V and U is also f-linear, then U can be extended to a subdivision Y_f as above. Therefore, it makes sense to refer to any such U as a linear piece for f.

We may describe the action of f on an f-linear piece U by a 4-tuple (A, V_i, V_j, U), where V_i and V_j are cells of Y, the domain U is an integral linear cone contained in V_i, the image $f(U)$ is contained in V_j, and A is the matrix describing the action of f on the linear piece U in the given coordinates of V_i and V_j.

In our algorithms, it will suffice to record only the triple (A, V_i, V_j), omitting the domain. Indeed, by multiplying the matrices from these triples, we can obtain a matrix for the product of mapping classes without knowing the domain. In the case where V_i is equal to V_j, we may write (A, V_i) instead of (A, V_i, V_j).

To illustrate, we return to our running example Y_Δ. Let

$$V_{1,1} = \{(x, y) \in V_1 \mid y \leq x\}, \text{ and}$$
$$V_{1,2} = \{(x, y) \in V_1 \mid x \leq y\}.$$

There is an integral cone complex map f_1 given by the four 4-tuples

$$\big(A_{1,1}, V_1, V_1, V_{1,1}\big), \big(A_{1,2}, V_1, V_2, V_{1,2}\big), (A_2, V_2, V_3, V_2), \text{ and } (A_3, V_3, V_3, V_3),$$

where

$$A_{1,1} = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right), \quad A_{1,2} = \left(\begin{smallmatrix} 1 & 0 \\ -1 & 1 \end{smallmatrix}\right), \quad A_2 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right), \quad A_3 = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right),$$

There is another cone complex map r given by permuting the V_1 cyclically, mapping each to the next by the identity matrix.

(b) If Y is an integral cone complex, there is a projectivized version of Y, denoted $\mathcal{P}(Y)$, which is a cell complex in the usual sense. For instance, the projectivized complex for $\mathrm{MF}(S)$ is $\mathcal{P}(\mathrm{MF}(S)) = \mathrm{PMF}(S)$ (Thurston described the $\mathrm{Mod}(S)$-action on $\mathrm{PMF}(S)$ as opposed to $\mathrm{MF}(S)$). If $f : Y \to Y$ is an integral cone complex map, then there is an induced cellular map $\mathcal{P}(Y_1) \to \mathcal{P}(Y_1)$, where Y_1 is an integral cone complex subdivision of Y. We also define a vertex ray of an integral cone complex Y to be the ray corresponding to a vertex of $\mathcal{P}(Y)$.

For our running example Y_Δ, the cell complex $\mathcal{P}(Y_\Delta)$ is a triangle. Under the geometric realization given by Figure 1, $\mathcal{P}(Y_\Delta)$ is identified with the triangle shown in the picture. There are three vertex rays in Y_Δ. The subdivision associated to the map f_1 is given by adding a vertex to $\mathcal{P}(Y_\Delta)$ corresponding to the ray spanned by $(1, 1)$ in V_1; this is the midpoint of the V_1-edge of $\mathcal{P}(Y_\Delta)$.

(c) We now set out to define the radius of an integral cone complex Y – which we also refer to as the radius of $\mathcal{P}(Y)$. First, we define a graph $\mathcal{G}_Y$ whose vertices are the top-dimensional cells of $\mathcal{P}(Y)$ and whose edges connect cells of $\mathcal{P}(Y)$ with nonempty intersection. For a set of vertices W of $\mathcal{G}_Y$, we define the radius of Y at W to be the supremal distance in $\mathcal{G}_Y$ from a vertex of $\mathcal{G}_Y$ to W. For a vertex c of $\mathcal{P}(Y)$, we define the radius of Y at c to be the radius of Y at W_c, the set of top-dimensional cells of $\mathcal{P}(Y)$ containing c.

For the running example Y_Δ, the graph $\mathcal{G}_Y$ is a triangle (dual to $\mathcal{P}(Y_\Delta)$), and so the radius with respect to any W is at most 1. As another example, if Y is any triangulation of S^2 given by the equator and any number of longitudes, then the radius of Y with respect to the vertex at the north pole is 1. As we will see, this example mirrors the situation for $\mathrm{MF}(S)$.

(d) Suppose that G is a group that acts on the integral cone complex Y by integral cone complex maps, and say that G is generated by $\Gamma = \{h_\ell\}$. We denote by Y_Γ the collection of integral cone complex subdivisions $\{Y_{h_\ell}\}$ (so there is one subdivision for each generator). As above, we denote the cells of the h_ℓ-subdivision of V_i by $V_{i,k}^{h_\ell}$.

If Γ is finite, then there is a single subdivision (namely, the common refinement of the Y_{h_ℓ}) with the property that each cell of the subdivision is a linear piece for each h_ℓ and h_ℓ^{-1}. In practice, it is convenient to use a different subdivision for each generator since this reduces the number of calculations required.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

(e) Let Y be an integral cone complex and let $f : Y \to Y$ be an integral cone complex map. For a point v in Y and a $\lambda \geq 0$, the point λv is a well-defined point of Y. With this in mind, a PL-eigenvector for f is a point $v \in Y$ so that $f(v) = \lambda v$ for some $\lambda \geq 0$. We emphasize that if A represents f on a cone V of Y, a real eigenvector v for A is not necessarily a PL-eigenvector for f. This is because v might not lie in the cone of $\mathbb{R}^d$ corresponding to V.

In our running example, the point $(1, 0) \in V_1$ is a PL-eigenvector of the integral cone complex map f_1 described above, since it is an eigenvector of the matrix $A_{1,1}$ and since the domain of this linear piece is contained in the codomain. By the same token, $(0, 1) \in V_1$ is not a PL-eigenvector of f_1: it is an eigenvector of $A_{1,2}$ but its image is $(0, 1) \in V_2$, which is not identified with (any multiple of) $(0, 1) \in V_1$.

(f) Next, a PL-eigenregion for the integral cone complex map f is a linear piece U for f that contains a PL-eigenvector v. The linear transformation associated to $f|U$ will recognize v as an eigenvector. More specifically, if $f|U$ has the triple (A, V_i, V_j) with $i = j$ then v will be a (classical) eigenvector for A. When $i \neq j$, we can choose coordinates on V_i and V_j that agree on the cell $V_{ij} = V_i \cap V_j$. In these coordinates, we again have that v is an eigenvector for A.

As part of our algorithm, we will find a point w with the property that it is contained in the interior of a union of PL-eigenregions for v. In such a case, there may be several different triples associated to the action of f at w, and hence several different matrices. However, all of these matrices will have v as an eigenvector.

In the running example, the map f_1 has two PL-eigenregions for the PL-eigenvector $(1, 0) \in V_1$, namely, $V_{1,1}$ and V_3. We already discussed the former region. Similarly, V_3 is a PL-eigenregion for f_1 since $(0, 1)$ is an eigenvector for A_3 and the third piece of f_1 has the same domain and codomain, which contains $(0, 1)$.

(g) Next, an integral cone complex map $f : Y \to Y$ is projectively source-sink if the induced cellular map $\mathcal{P}(Y) \to \mathcal{P}(Y)$ has source-sink dynamics. This means that the induced map has exactly two fixed points, the source and the sink, and that the f-iterates of all points outside of the source converge to the sink. In particular, the only two PL-eigenvectors for f (up to scale) are representatives of the source and the sink.

In the running example, if we take $f_2 = r \circ f_1 \circ r^{-1}$, then the integral cone complex map

$$f_2^{-1} \circ f_1$$

is projectively source-sink. The sink and source correspond to the vectors $(\varphi, 1) \in V_3$ and $(1, \varphi) \in V_1$, where φ is the golden mean. The corresponding triples are

$$(D, V_3, V_3) \text{ and } (E, V_1, V_1)$$

where

$$D = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right) \text{ and } E = \left(\begin{smallmatrix} 2 & -1 \\ -1 & 1 \end{smallmatrix}\right).$$

(h) Let Y be an integral cone complex and $f : Y \to Y$ a projectively source-sink integral cone complex map. Let U be a PL-eigenregion for the sink of f with triple (A, V_i, V_j). There exists a maximal subcone W of $V_i \cap V_j$ invariant under f, and the map $f|W$ is a dynamical map in the sense that f maps W to itself. The cone W is nontrivial since it contains the ray corresponding to the sink. In Section 1.2 we give an algorithmic procedure to find W and to compute the linear map associated to $f|W$. We also show that, in any $\mathbb{R}^d$-coordinates on V_i, the cone W is the intersection of V_i with a subspace of $\mathbb{R}^d$. Hence it makes sense to project V_i onto W.

We prove in Lemma 7.4 that (under a certain condition about the location of the sink) the largest real eigenvalue of $f|W$ is the eigenvalue for the sink, and any corresponding positive eigenvector is the sink. If we precompose $f|W$ with a projection map $V_i \to W$ (really a projection of the corresponding subspaces of $\mathbb{R}^d$), we obtain a linear map $V_i \to W$, which we call an extended dynamical map for the sink; say that D is a matrix for this map. The nonzero real eigenvalues of D are the same as those

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

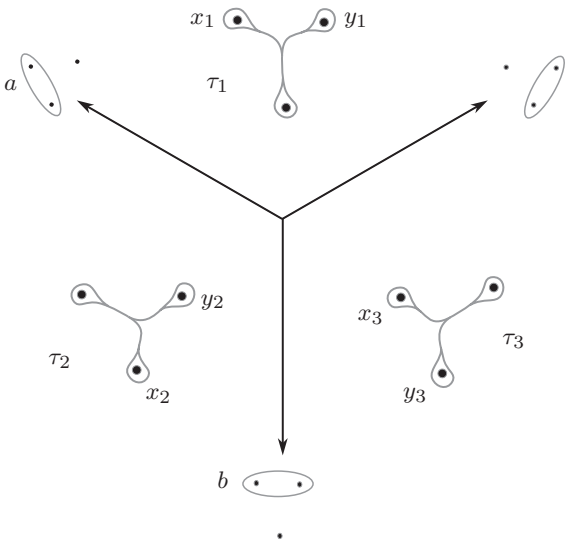


Figure 2. *Train tracks giving a cell structure on* $\mathrm{MF}(S_{0,4})$.

of $f|W$. As such, the pair (V_i, D) contains the data of the sink; we refer to any such pair as a sink package.

In the running example, we have $i = j$ (this is the generic scenario); specifically $W = V_i = V_j = V_3$. As such, the sink package for the map $f_2^{-1} \circ f_1$ is (V_3, D), where D is the matrix given above.

The space of measured foliations.... Turning back now to the mapping class group, we denote by $\mathrm{MF}(S)$ the space of measured foliations on S, again up to Whitehead moves and isotopy. See the book by Fathi–Laudenbach–Poénaru [12] for background on $\mathrm{MF}(S)$.

Let $S = S_{g,p}$ and let $\xi(S) = 3g - 3 + p$. Thurston proved that $\mathrm{MF}(S)$ is homeomorphic to $\mathbb{R}^{2\xi(S)}$. He further proved that $\mathrm{MF}(S)$ has the structure of an integral cone complex [30]. Thurston actually described the projectivized complex for $\mathrm{PMF}(S)$, which can be described in terms of train track coordinates as follows.

A train track in S is a certain type of embedded finite graph in S with a well-defined tangent line at each point; see the book of Penner–Harer [25] for details. Thurston's cells in the cell structure for $\mathrm{PMF}(S)$ – equivalently, the projectivized integral cone complex structure for $\mathrm{MF}(S)$ – correspond naturally to train tracks in S. Conversely, every train track gives a cone in $\mathrm{MF}(S)$ by considering the measured foliations that are carried by τ. Given a train track τ, the associated cone will be denoted $V(\tau)$. For any given S, we may choose once and for all a cell decomposition Y of $\mathrm{MF}(S)$ with cells $V(\tau_1), \ldots, V(\tau_N)$.

There are infinitely many choices for the set of $V(\tau_i)$, that is, infinitely many choices of structures as an integral cone complex where the cells correspond to train tracks. But they all give the same PL structure. While there are integral cone complex structures on $\mathrm{MF}(S)$ that do not arise from train tracks, we will only make use of the ones that do; in what follows we refer to these as *train track cell structures*.

As an example, let S be $S_{0,4}$, or equivalently the thrice-punctured plane. The train tracks τ_1, τ_2, and τ_3 shown in Figure 2 give a cell structure on $\mathrm{MF}(S_{0,4})$. The weights x_i and y_i on τ_i give an identification of each cell $V(\tau_i)$ with the first quadrant in $\mathbb{R}^2$. The resulting integral cone complex structure on $\mathrm{MF}(S_{0,4})$ is isomorphic to our running example Y_Δ. The geometric realization given in Figure 1 has a geometric description, namely, it is given by the geometric intersection numbers of a weighted train track with three rays to infinity that emanate from the three marked points.

There is a convenient class of train track cell structures on $\mathrm{MF}(S)$ that applies to an arbitrary S, namely, the ones arising from Dehn–Thurston train tracks. These are defined as follows. First, we choose a pants decomposition of S; we will refer to the curves of the pants decomposition as base curves. From there, we decompose S into a collection of annuli and pairs of pants. The core curves of the annuli are the base curves and the pairs of pants are the closures of the complementary regions. Each pair of pants

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

has between one and three boundary components, with the other holes accounted for by punctures. As in the book by Penner–Harer [25, §2.6], there are four model (sub)-train tracks on each annulus and at most four on each pair of pants. By gluing together all possible models, we obtain the collection of Dehn–Thurston train tracks. We prove in Lemma 7.3 that if Y is a Dehn–Thurston cell structure on $\mathrm{MF}(S)$ and c is one of the base curves, then the radius of Y at c is 1.

...and the mapping class group action. There is a natural action of $\mathrm{Mod}(S)$ on $\mathrm{MF}(S)$; again see [12]. As explained by Thurston, the action of $\mathrm{Mod}(S)$ on $\mathrm{MF}(S)$ is by integral cone complex maps (in his setup, piecewise integral projective maps).

Given an integral cone complex structure Y for $\mathrm{MF}(S)$ and a finite generating set Γ for $\mathrm{Mod}(S)$, there is (as above) an associated collection of subdivisions Y_Γ (as in our general discussion, there are canonical choices for these subdivisions, but in practice we compute subdivisions which may be finer than those). Moreover, each subdivision is a train track cell structure. If $V(\tau_i)$ is a top-dimensional cell of Y and h_ℓ is an element of Γ, we denote by $V(\tau_{i,1}^{h_\ell}), \ldots, V(\tau_{i,N}^{h_\ell})$ the cells of the corresponding subdivision of $V(\tau_i)$. Each $\tau_{i,j}^{h_\ell}$ is obtained from the corresponding τ_i by a sequence of splittings; again see [25] for details.

An additional property of the $\mathrm{Mod}(S)$-action is that each pseudo-Anosov element acts on $\mathrm{MF}(S)$ with projective source-sink dynamics. In particular, such an element has exactly two PL-eigenvectors. The eigenvalues are λ and $1/\lambda$, for the source and sink, respectively. We will refer to a PL-eigenregion for either of these PL-eigenvectors as a Nielsen–Thurston eigenregion.

From a Nielsen–Thurston eigenregion, we obtain an extended dynamical map, and a corresponding sink package $(V(\tau_i), D)$, or simply (τ_i, D). Any such sink package (τ_i, D) will be called a Nielsen–Thurston package. This is the desired output of our algorithm.

In our running example, the left-handed half-twist H_a about the curve a in Figure 2 (obtained by a clockwise π-rotation of the two leftmost marked points) is given by the map f_1 described above (using the above identification of $\mathrm{MF}(S_{0,4})$ with Y_Δ), and similarly the half-twist H_b is given by f_2. The mapping class $H_b^{-1}H_a$ is pseudo-Anosov with stretch factor φ^2. By the above calculation, the sink package for $H_b^{-1}H_a$ is $(D, V(\tau_3))$.

Statement of the main technical theorem. Let $S = S_{g,p}$. For the statement of Theorem 1.1, we must define a constant $Q = Q(S)$. We build up the definition of Q by defining a sequence of preliminary constants. We do this in order to give a sense of how Q arises, and to make it feasible to improve our Q by improving the intermediate constants (and hence improve the implementation). In the end, we give a definition of Q using the intermediary constants, and also a formula given directly in terms of g and p; the impatient reader is encouraged to skip to the latter definition of Q.

We begin with a basic constant related to the topology of S:

$$K(S) = 4\xi(S) = 4(3g - 3 + p).$$

The number $K = K(S)$ is the maximum number of separatrices for a measured foliation on S; we refer to K as the separatrix complexity of S. In other words, this is the maximum possible sum of the degrees of the singularities of a foliation on S.

Next we have the constants

$$D(S) = 1 \quad \text{and} \quad E(S) = 14.$$

The number $D = D(S)$ is the radius (as defined above) of any Dehn–Thurston integral cone complex structure on $\mathrm{MF}(S)$, specifically the radius about one of the curves of the underlying pants decomposition; see Lemma 7.3. The number $E(S)$ is an upper bound on the diameter in $\mathcal{C}(S)$ of the set of vertex cycles for a train track in S. Aougab [2, Section 5] proved that for all but finitely many S there is an upper bound of 3 for all train tracks in S. As observed by Tang–Webb [29, Section 2.3], it follows from the work of Bowditch [8] that there is an upper bound of 14 for all train tracks in all surfaces. Masur–Minsky [20] had previously conjectured that the diameter is bounded above by 3 for all train tracks in all surfaces, but this is still an open problem.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

From the above constants, we now define several constants that appear in the statements of the main propositions in the paper or in the proof of Theorem 1.1: Next, we have

$$\begin{aligned} Q_{\text{curve}}(S) &= 2K^2 \\ Q_{\text{track}}(S) &= E \cdot Q_{\text{curve}} \\ Q_{\text{force}}(S) &= Q_{\text{track}} + (2K+6) \cdot 2K \\ Q_{\text{fit}}(S) &= Q_{\text{force}}(S) + (E+3) \cdot Q_{\text{curve}} \\ Q_{\text{DT}}(S) &= (D+1)Q_{\text{track}} \end{aligned}$$

The first three of these constants appear in the statements of Propositions 3.1, 5.1, and 6.1, which we call the bounded slope lemma, the forcing lemma, and the fitting lemma. The number $2K + 6$ appears in the statement of Lemma 4.1.

Finally, we define

$$Q(S) = Q_{\text{force}}(S) + Q_{\text{fit}}(S) + Q_{\text{DT}}(S) + 1.$$

This gives

$$\begin{aligned} Q(S_{g,p}) &= 2,464\,\xi^2 + 96\,\xi + 1 \\ &= 2,464(3g-3+p)^2 + 96(3g-3+p) + 1 \end{aligned}$$

and in particular

$$Q(S_2) = 22,465 \quad \text{and} \quad Q(S_3) = 89,281.$$

The coefficient 2,464 arises as $32(5E + 7)$. If one proves the Masur–Minsky conjecture that $E(S)$ can be replaced with 3, then we obtain

$$Q(S_{g,p}) = 704\,\xi^2 + 96\,\xi + 1, \quad Q(S_2) = 6,625, \quad \text{and} \quad Q(S_3) = 25,921.$$

We may now state our main technical result in terms of Dehn–Thurston cell structures, Nielsen–Thurston eigenregions, and the number $Q = Q(S)$ just defined.

Theorem 1.1. *Fix a surface $S = S_{g,p}$, a generating set Γ for $\mathrm{Mod}(S)$, a Dehn–Thurston cell structure Y of $\mathrm{MF}(S)$, and a base curve $c \in \mathrm{MF}(S)$ for the Dehn–Thurston coordinates. Let $Q = Q(S)$. Let $f \in \mathrm{Mod}(S)$ be pseudo-Anosov with unstable foliation $\mathcal{F}$ and stretch factor λ. Then*

1. *$f^Q(c)$ lies in the interior of a union of Nielsen–Thurston eigenregions for $\mathcal{F}$, and*
2. *if $(V(\tau_i), D)$ is a Nielsen–Thurston package for f, then λ is the largest real eigenvalue of D and any corresponding positive eigenvector is a positive multiple of $\mathcal{F}$.*

We may think of the first statement of Theorem 1.1 as saying that the Nielsen–Thurston eigenregion is large in some sense (Proposition 6.1 even more directly has this interpretation). For a general PL-action of a group on a manifold, we would expect the eigenregion for a sink to become arbitrarily small as the word length of f becomes increasingly large. As such, Theorem 1.1 is not a general fact about PL-actions; our arguments rely heavily on technology specific to the setting of the mapping class group.

The point of the first statement of Theorem 1.1 is that $f^Q(c)$ finds a Nielsen–Thurston eigenregion for f (or perhaps several). From the action of f on any such eigenregion we may compute a Nielsen–Thurston package (τ, D). Then the second statement of the theorem allows us to extract the stretch factor and unstable measured foliation for f.

We emphasize that the cell structure on $\mathrm{MF}(S)$ is fixed once and for all. If the cell structure – and hence the notion of linearity – were allowed to depend on f, then the corresponding statement would hold for $Q = 1$, because there exists an invariant train track for f that carries the curve c. We also emphasize

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

that our exponent Q does not depend on the generating set for $\mathrm{Mod}(S)$. We will return to this point in Section 1.2.

1.2. The algorithm

We now explain in detail the algorithm that comes from Theorem 1.1. As above, the algorithm runs in quadratic time with respect to the length of the spelling of f as a word in any fixed generating set for $\mathrm{Mod}(S)$, and it runs in polynomial time with respect to $\xi(S)$. And the output is a train track and a matrix. We state this result as Corollary 1.2 below.

There are three subroutines in the main algorithm: the linear piece algorithm, the basic computation algorithm, and the sink package algorithm. We explain these in turn. The first of these is used in the second; the second and third are both utilized in the main algorithm.

Ledgers and the linear piece algorithm. Let M be a space that admits an integral cone complex structure Y, and let G be a finitely generated group acting on M by integral cone complex maps. Suppose we have chosen

1. a generating set $\Gamma = \{h_1, \dots, h_m\}$ for G and
2. a cell structure for Y on M with cells $\{V_1, \dots, V_N\}$.

A linear piece algorithm for the pair (Y, Γ) is an algorithm that takes as input an $h_\ell \in \Gamma$ and a cell V_i of Y, and produces as output two pieces of data:

3. a subdivision $\{V_{i,k}^{h_\ell}\}$ of V_i corresponding to h_ℓ and
4. the integer matrices $\mathcal{A} = \{A_{i,k}^{h_\ell}\}$ so each $h_\ell | V_{i,k}^{h_\ell}$ is described by $A_{i,k}^{h_\ell}$.

Here the $V_{i,k}^{h_\ell}$ are described in terms of linear inequalities in the V_i-coordinates.

In some applications it may be expedient to write down once and for all the set of subdivisions $Y_\Gamma = \{Y_{h_\ell}\}$ for all cells V_i and all generators $h_\ell \in \Gamma$, as well as the set of all matrices $\mathcal{A} = \{A_{i,k}^{h_\ell}\}$. We would refer to any such quadruple $(\Gamma, Y, Y_\Gamma, \mathcal{A})$ as a *ledger* for the action of G on M. As above, once Γ and Y are chosen, the Y_Γ and the $\mathcal{A}$ are derived from those choices. In Section 2 we demonstrate our algorithm through a particular example, and in that case we start by giving the entire ledger.

Linear piece algorithms for mapping class groups. For mapping class groups, there are many examples of linear piece algorithms (and ledgers) already in the literature. In his thesis [27], Penner gave formulas – which give a constant-time linear piece algorithm – in terms of Dehn–Thurston train track coordinates (explicit matrices are not given, but they can be derived from the computations there). Penner's formulas contain minor mistakes, corrected in his subsequent paper [24]. There is still one other mistake, pointed out by Yandi Wu: in the second elementary move (Theorem A.2), the λ's and κ's should be switched [32]. For punctured surfaces, a constant-time linear piece algorithm is given by Hamidi-Tehrani–Chen [18].

Again, for the case of the $\mathrm{Mod}(S)$-action on $\mathrm{MF}(S)$, the cells of Y_Γ correspond to train tracks. Further, we can choose the ledger so that each subdivision $V_{i,k}^{h_\ell}$ is equal to some $V(\tau_{i,k}^{h_\ell})$, where $\tau_{i,k}^{h_\ell}$ is obtained from train track τ_i by a sequence of splittings; see the book by Penner–Harer for the definition of a splitting [25].

In the case of $\mathrm{Mod}(S)$, the number of cells V_i is exponential in $\xi(S)$. Therefore, in order to obtain an algorithm that runs in polynomial time with respect to $\xi(S)$, we cannot compute the entire ledger. Fortunately, this is not necessary. Penner's ledger (or formulas) gives us a linear piece algorithm that runs in constant time with the Lickorish generators for $\mathrm{Mod}(S)$. This is because (a) each generator only interacts with a bounded number of coordinates, (b) there is a finite set of closed formulas describing the action of a generator on the corresponding coordinates, and (c) the corresponding linear inequalities have bounded bit size.

Basic computation algorithm. Say that G is a group with finite generating set $\Gamma = \{h_i\}$ and that G acts on a space M that admits the structure of an m-dimensional integral cone complex Y. Suppose we

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

are given some word

$$w = f_n \cdots f_1$$

in the generating set Γ, and let $c \in M$ be a point, written as an integer point in some cell V_i of Y. Say that w represents $f \in G$. We would like to compute:

- $f(c) \in M$ and
- a triple (A, V_i, V_j) describing f in a partial neighborhood of c.

By the latter, we mean (as above), a triple (A, V_i, V_j) where V_i and V_j are cells of Y and A is a matrix describing f on a linear piece $U \subseteq V_i$ containing c (the matrix A uses the V_i and V_j coordinates). We now give the algorithm for performing these computations.

Setup. A space M and group G as above, with a choice of linear piece algorithm for the choices of Y and Γ.

Input. A word $w = f_n \cdots f_1$ in the generating set Γ (so each f_ℓ is some $h_{i_\ell}^{\pm 1}$)

Initialization. Let $c_0 = c$, with $c_0 \in V_{i_0}$, let $A_0 = I_m$ (the identity matrix), and let $\ell = 1$.

For $\ell = 1, \ldots, n$:

1. Apply the linear piece algorithm to $V_{i_{\ell-1}}$ and f_ℓ, and determine a top-dimensional subdivision cell $V^{f_\ell}_{i_{\ell-1},k}$ that contains $c_{\ell-1}$ and on which f_ℓ is linear, has codomain in some V_{i_ℓ}, and is represented by a matrix $A^{f_\ell}_{i,k}$.
2. Compute $c_\ell = A^{f_\ell}_{i,k} c_{\ell-1}$ and $A_\ell = A^{f_\ell}_{i,k} A_{\ell-1}$.

Output. $f(c) = c_n$ and (A_n, V_{i_0}, V_{i_n})

We emphasize that there can be different outputs for the triple (A_n, V_{i_0}, V_{i_n}), since at each stage there may be more than one linear piece of f_ℓ containing c. We also point out that step (1) is where the generating set for G is most relevant to the complexity of the algorithm; if we choose a generating set that is not well-suited to the cell structure on Y, then the linear piece algorithm can be more complicated.

The sink package algorithm. We now give an algorithm to find the sink package for a projectively source-sink integral cone complex map $f : Y \to Y$. We suppose that we have in hand a triple (A, V_i, V_j) for the action of f on a PL-eigenregion for the sink $v \in Y$. Both V_i and V_j are cones in $\mathbb{R}^d$, and we may assume without loss of generality that whatever faces are shared in Y are also shared in $\mathbb{R}^d$. In this case we say that the coordinates on V_i and V_j are compatible.

Input. A triple (A, V_i, V_j) for a PL-eigenregion for the sink of a projectively source-sink integral cone complex map. We assume that the matrix A is written in terms of compatible $\mathbb{R}^d$-coordinates on V_i and V_j.

Initialization. Let W_0 be the subspace of $\mathbb{R}^d$ spanned by $V_i \cap V_j$ (in the compatible coordinates).

1. *For* $k = 1, \ldots, m$:
   let $W_k = A(W_{k-1}) \cap W_{k-1}$
2. Choose a $d \times d$ matrix B that projects $\mathbb{R}^d$ to W_m; set $D = AB$.

Output. (V_i, D)

In general, projection onto an integral linear subspace gives a matrix with rational entries. So, generically, we expect the output matrix D to have rational entries.

We will of course apply the sink package algorithm in the case of the mapping class group, in which case V_i can be replaced by the train track defining the cell of $\mathrm{MF}(S)$ corresponding to V_i. Again, we prove in Lemma 7.4 that the largest real eigenvalue of D is the eigenvalue for the sink, and any corresponding positive eigenvector is the sink.

Property Q. We will state the main algorithm for any group action on an integral cone complex that satisfies the same property as the one guaranteed for the mapping class group by the first statement of Theorem 1.1. Specifically, we say that the action of a finitely generated group G on an integral cone complex Y satisfies property Q if there is a natural number $Q = Q(Y)$ so that, for any $f \in G$ acting

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

with projectively source-sink dynamics and any $c \in Y$ lying on a vertex ray, the point $f^Q(c)$ lies in the interior of a union of PL-eigenregions for f containing the sink.

The main algorithm. Say that G is a group with finite generating set $\Gamma = \{h_i\}$ and that G acts on a space M that admits the structure of an m-dimensional integral cone complex Y. Fix a constant-time linear piece algorithm for the choices of Y and Γ, and fix some nonzero $c \in M$ lying on a vertex ray of Y. Suppose that the G-action on M satisfies property Q for some specific $Q = Q(Y)$.

Input. A word $w = f_n \cdots f_1$ in the generating set Γ that represents a projectively source-sink element f of G.

Step 1. Applying the basic computation algorithm to w, compute $f^Q(c)$.

Step 2. Applying the basic computation algorithm to w, compute a triple (A, τ_i, τ_j) associated to an f-linear piece U of Y containing $f^Q(c)$.

Step 3. Compute a sink package (τ_i, D) for f using $V(\tau_i)$-coordinates.

Output. (τ_i, D)

We are finally ready to state our corollary of Theorem 1.1, which gives the algorithm from the Main Theorem.

Corollary 1.2. *Fix a surface S and let Γ be any finite generating set for* $\mathrm{Mod}(S)$. *The main algorithm computes a Nielsen–Thurston package for any pseudo-Anosov $f \in \mathrm{Mod}(S)$, given as a word w in Γ. Assuming that there is a constant-time linear piece algorithm for the choice of Γ and the choice of integral cone complex structure on* $\mathrm{MF}(S)$, *the running time of the main algorithm is quadratic in the length of w and polynomial in $\xi(S)$.*

Corollary 1.2 follows immediately from Theorem 1.1, the standard running times for matrix multiplication and the Zassenhaus algorithm, and the existence of a constant-time linear piece algorithm for the mapping class group. (The Zassenhaus algorithm here boils down to row reduction of matrices of size $O(\xi)$ and bit size $O(n)$, and this row reduction problem has the complexity stated in the introduction; see [33, Lecture X, Theorem 5]).

The Nielsen–Thurston problem. The more general problem that our work fits into is the Nielsen–Thurston problem. This is the algorithmic problem of determining, from a product of generators of $\mathrm{Mod}(S)$, the Nielsen–Thurston type of the resulting $f \in \mathrm{Mod}(S)$ and the associated data: the period if f is periodic, the canonical reduction system if f is reducible, and the stretch factor and foliations when f is pseudo-Anosov.

Bell–Webb gave a polynomial-time algorithm for the reducibility problem, that is, the problem of determining whether an infinite order mapping class (given as a word in the generators) is reducible [5]. In this case, their algorithm gives the canonical reduction system. Otherwise, their algorithm computes the asymptotic translation length in the curve graph. (It is known classically that if f is periodic, then this – plus the period – can be determined in quadratic time. Indeed, this follows from the upper bound on the period (depending only on S) and the quadratic-time solution to the word problem for $\mathrm{Mod}(S)$). Thus, combining our work with theirs, we obtain a polynomial-time algorithm to solve the Nielsen–Thurston problem.

Baroni [3] recently gave an alternate version of the Bell–Webb algorithm. Instead of just being polynomial in the word length, his algorithm is polynomial in the word length and the complexity of the surface. On the other hand, Baroni's algorithm does not explicitly determine any reducing curves.

As pointed out to us by Agol, there is an approach to the Nielsen–Thurston problem using the classical theory of 3-manifolds. Given a mapping class, the algorithm proceeds by producing a triangulation of the corresponding mapping torus, and then using the Jaco–Tollefson algorithm to find the JSJ-decomposition [19, Section 8]. (The key point is that a pseudo-Anosov mapping class has a trivial JSJ-decomposition.)

1.3. New tools and an outline of the proof

We now give an outline of the paper and the proofs of the main results, as well as descriptions of the new tools introduced in the paper. We begin in Section 2 by illustrating our algorithm through a concrete example.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

The rest of the paper is devoted to the proof of Theorem 1.1. The keys to our proof are Propositions 5.1 and 6.1, called the forcing lemma and the fitting lemma. Both propositions are stated in terms of the notion of slope for a curve in a surface, which is also new. We now turn to developing this idea, beginning with the required background.

Horizontal and vertical foliations. Let $f \in \mathrm{Mod}(S)$ be a pseudo-Anosov mapping class. In order to orient ourselves (and the reader), we will refer to the stable and unstable measured foliations, $(\mathcal{F}^s, \mu^s)$ and $(\mathcal{F}^u, \mu^u)$, as $(\mathcal{F}_v, \mu_v)$ and $(\mathcal{F}_h, \mu_h)$. The 'v' and 'h' stand for "vertical" and "horizontal" and in our pictures we will (away from singularities) draw the foliations so that $\mathcal{F}_v$ is vertical, $\mathcal{F}_h$ is horizontal, and the transverse measures are $|dx|$ and $|dy|$. In such a picture, the pseudo-Anosov representative of f stretches horizontally and compresses vertically.

Singular Euclidean structures. The two measured foliations $\mathcal{F}_v$ and $\mathcal{F}_h$ induce a singular Euclidean structure X on S, the singular points being the singularities of the foliations. Let $\bar{X}$ be the closed singular Euclidean surface defined by taking the completion of X with the induced metric. There is a set of marked points $\Sigma \subseteq \bar{X}$ so that $\bar{X} \setminus \Sigma$ is equal to the surface obtained from X by deleting the singularities of $\mathcal{F}_v$ and $\mathcal{F}_h$. In other words, Σ is the set of points that either correspond to punctures of X or singularities of the foliations (or both).

We remark that $\mathcal{F}_h$ and $\mathcal{F}_v$ are also the horizontal and vertical measured foliations associated to an integrable meromorphic quadratic differential q on $\bar{X}$, regarded as a closed Riemann surface.

Section 3: Saddle connections and slope. Let X be a singular Euclidean surface. A saddle connection in the surface $\bar{X}$ is a closed geodesic arc that meets Σ in exactly two points, its endpoints. The corresponding arc in X is also called a saddle connection.

Let σ be such a saddle connection in X. We define the *slope* of σ as its height over width:

$$\mathrm{slope}_X(\sigma) = \frac{\mu_h(\sigma)}{\mu_v(\sigma)}.$$

For any collection of saddle connections Z, we define $\mathrm{slope}_X(Z)$ to be the set of slopes of elements of Z.

Let c be a simple closed curve in X. By pulling c tight in X, the curve c degenerates to an immersed path in $\bar{X}$ contained in a set of pairwise noncrossing saddle connections. Here and throughout when we say that saddle connections are noncrossing, we mean that their interiors are disjoint. Similarly, if c is a measured foliation on S, we can pull tight each nonsingular leaf. Each leaf gives an immersed path contained in a set of pairwise noncrossing saddle connections. Further, since the leaves are pairwise disjoint, there is a single collection of pairwise noncrossing saddle connections so that each leaf gives an immersed path contained in the union. For c either a simple closed curve or a foliation, let $Z = \{\sigma_1, \ldots, \sigma_k\}$ be the resulting collection of pairwise noncrossing saddle connections. We define the slope of c as the finite set

$$\mathrm{slope}_X(c) = \mathrm{slope}_X(Z) = \{\mathrm{slope}_X(\sigma_1), \ldots, \mathrm{slope}_X(\sigma_k)\}.$$

Finally, for a train track τ, we define $\mathrm{slope}_X(\tau)$ to be the union of the slopes of the vertex cycles.

Section 3 (cont.): Horizontality. It is sometimes convenient for us to use an alternate formulation of the slope. Let $f : S \to S$ be a pseudo-Anosov mapping class with stretch factor λ and singular Euclidean surface X, and let c (or τ) be as above. We define

$$\mathcal{H}_f(c) = -\frac{1}{2}\log_\lambda \mathrm{slope}_X(c).$$

We refer to $\mathcal{H}_f(c)$ as the horizontality of c. We are using an abuse of notation here, as $\mathcal{H}_f$ is not completely determined by f. Indeed, different choices of X lead to horizontality functions that differ by a constant. For any such choice, horizontality satisfies the equation

$$\mathcal{H}_f(f^k(c)) = \mathcal{H}_f(c) + k$$

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

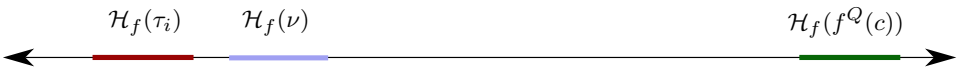


Figure 3. *In the proof of Theorem 1.1, we choose Q so $\mathcal{H}_f(f^Q(c))$ is very large, we find a coordinate train track τ_i carrying $\mathcal{F}_h$, and we choose an f-invariant ν carried by τ_i with $\mathcal{H}_f(\nu)$ small.*

for any k. We also prove in Proposition 3.1 that for any set of disjoint curves, arcs, or leaves of a foliation c, the set $\mathcal{H}_f(c)$ has diameter bounded above by $Q_{\text{curve}} = 2K^2$, the key point being that this bound depends only on S (and not on f).

For two sets of real numbers A and B we say that $A < B$ if $\max A < \min B$. We will use this notion to compare the slopes and horizontalities of the various objects discussed above.

As an application of our upper bound on $\operatorname{diam} \mathcal{H}_f(c)$, we give a new approach to, and – in the case of a closed surface – a slight improvement of the result of Gadre–Tsai on stable translation distances in the curve graph; see Section 3.4.

Section 4: Approximating rectangles. In Section 4 we prove Lemma 4.1, a technical lemma that will be used to prove the forcing lemma in Section 5. The basic idea of the lemma is to approximate a configuration $(\gamma_1, \gamma_2, \ell)$, where the γ_i are crossing saddle connections and ℓ is a horizontal leaf, by a configuration (γ_1, γ_2, R), where R is a maximal rectangle. For the latter to be an approximation, we mean that the γ_i cross the horizontal sides of R and that the slope of R – defined to be the slope of either diagonal – is not too small.

Section 5: The forcing lemma. Let f, $\mathcal{F}_h$, and X be as above, and let τ be a filling, transversely recurrent train track in S. We write $c \prec \tau$ and $\mathcal{F} \prec \tau$ to mean that a curve c or a foliation $\mathcal{F}$ is carried by τ, and in this case we write $\tau[c]$ and $\tau[\mathcal{F}]$ for the sub-train track of τ consisting of all branches traversed by c or $\mathcal{F}$, respectively.

Suppose that c is a simple closed curve or a leaf of a measured foliation in S. In the forcing lemma (Proposition 5.1) we assume that

$$\mathcal{H}_f(c) \geq \mathcal{H}_f(\tau) + Q_{\text{force}}.$$

Under this assumption, the forcing lemma gives the following implications:

1. If $c \prec \tau$, then $\mathcal{F}_h \prec \tau[c]$.
2. If $\mathcal{F}_h \prec \tau$, then c is carried by a diagonal extension of $\tau[\mathcal{F}_h]$.

Again, for the relevant train track-related definitions, including carrying and diagonal extensions, see the book by Penner–Harer [25].

Section 6: The fitting lemma. Suppose f is a pseudo-Anosov mapping class and that τ_1 and τ_2 are two train tracks carrying the horizontal foliation $\mathcal{F}_h$. The fitting lemma (Proposition 6.1) gives an f-invariant train track ν – meaning that $f(\nu) \prec \nu$ – with two properties:

1. $\nu \prec \tau_i$ for $i \in \{1, 2\}$ and
2. $\mathcal{H}_f(\nu) \leq \mathcal{H}_f(\tau_i) + Q_{\text{fit}}$ for $i \in \{1, 2\}$.

The train tracks τ_1 and τ_2 will be Dehn–Thurston train tracks coming from the cell structure on $\mathrm{MF}(S)$. On an intuitive level, if a train track ν is carried by τ_i and has much larger horizontality, then $V(\nu)$ is a very small subregion of $V(\tau_i)$. So the ν in the conclusion of Proposition 6.1 is fitted in the sense that $V(\nu)$ is both contained in $V(\tau_i)$ and is large (intuitively, we think of $V(\nu)$ as being almost the size of $V(\tau_i)$, although it may be of smaller dimension).

Section 7: Proof of the main technical theorem. We sketch the idea of the proof of Theorem 1.1. Choose any curve c lying on a vertex ray of $\mathrm{MF}(S)$. Our Q and Q_{DT} are defined exactly so that

$$\mathcal{H}_f(f^Q(c)) \geq \mathcal{H}_f(\tau_i) + Q_{\text{force}} + Q_{\text{fit}} + 1$$

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

for all train tracks τ_i arising in the chosen cell decomposition of $\mathrm{MF}(S)$. Suppose that $f^Q(c)$ is carried by the coordinate train track τ_i. By the forcing lemma, $\mathcal{F}_h$ is carried by τ_i. By the fitting lemma, there is an f-invariant train track ν so that

$$\mathcal{H}_f(f^Q(c)) \geq \mathcal{H}_f(\nu) + Q_{\text{force}} + 1$$

Then for every curve d near $f^Q(c)$, the forcing lemma (plus the continuity of horizontality) gives a diagonal extension ν_d of ν that carries d. Each such $V(\nu_d)$ is a Nielsen–Thurston eigenregion. Since the $V(\nu_d)$ form an open neighborhood of $f^Q(c)$ in $\mathrm{MF}(S)$, every linear piece for f that contains $f^Q(c)$ must overlap one of the Nielsen–Thurston eigenregions $V(\nu_d)$, and hence is contained in a Nielsen–Thurston eigenregion, as desired.

2. An example

In this section, we apply (a version of) our algorithm to a specific example of a pseudo-Anosov mapping class. Our example lies in the braid group B_3. This group is naturally isomorphic to the mapping class group of D_3, the closed disk with three marked points in the interior. It is also naturally isomorphic to the mapping class group of the surface obtained from a torus by deleting the interior of a closed, embedded disk.

The standard generating set for B_3 is $\Gamma = \{\sigma_1, \sigma_2\}$. Each generator σ_i is a half-twist. We consider the pseudo-Anosov braid

$$f = \sigma_2\sigma_1^{-1}\sigma_2\sigma_1\sigma_1\sigma_1$$

We will use our algorithm to show that the stretch factor of f is

$$\lambda = \frac{5 + \sqrt{21}}{2}$$

and to obtain a description of the unstable foliation for f in terms of train tracks.

Our method in this section diverges from the main algorithm of this paper in that we do not make use of the explicit Q from Theorem 1.1. Instead, we use a guess-and-check approach. Specifically, we use the basic computation algorithm to compute $f(c)$ and a triple (A, V_i, V_j) for the action of f at $f(c)$. Then we find an eigenvector v for A and check directly that it is a PL-eigenvector for f with eigenvalue $\lambda > 1$; it follows that v represents the sink (if v were not a PL-eigenvector, we would proceed by computing the action of f at $f^2(c)$, etc.). This method always terminates, and in practice it can be more efficient than our main algorithm (since Q is so large). Our Theorem 1.1 guarantees that this method will terminate in quadratic time.

The discussion in this section has one other small difference from our general setup, namely, the fact that D_3 has boundary; on the other hand, we could delete the boundary and consider f as an element of $\mathrm{Mod}(S_{0,4})$ without changing the stretch factor or foliations. Note that

$$Q(S_{0,4}) = 2,464 + 96 + 1 = 2,561$$

which is much larger than the power of f required for the guess-and-check method, which in this case is 1.

2.1. The ledger

We start by setting up the ledger for the B_3-action on $\mathrm{MF}(\mathrm{D}_3) \cong \mathbb{R}^2$. Since we already defined the generating set Γ, it remains to choose a cell decomposition Y for $\mathrm{MF}(\mathrm{D}_3)$ and then find the collection of subdivisions Y_Γ and the resulting collection of matrices $\mathcal{A}$.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

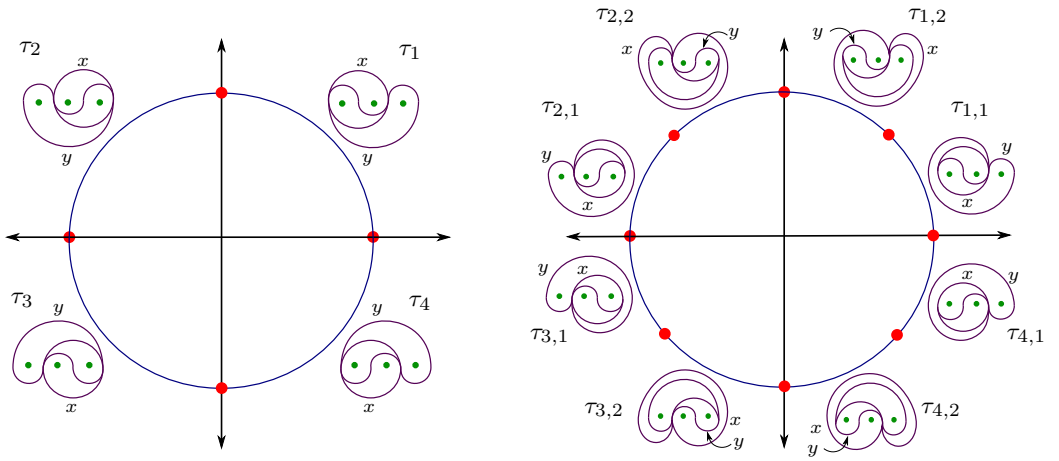


Figure 4. *The integral cone complex structure Y on* $\mathrm{MF}(\mathrm{D}_3)$ *and the common refinement of* Y_Γ.

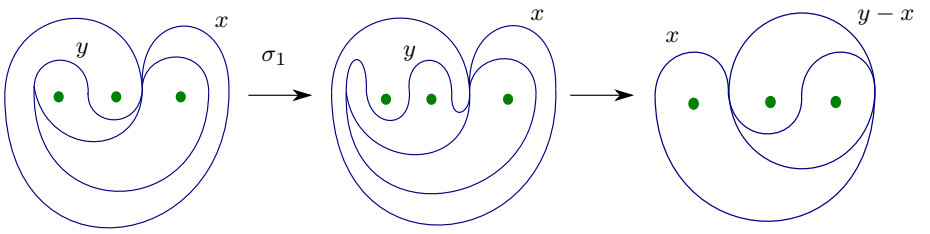


Figure 5. *A part of the computation of the linear transformation* $V_{12} \to V_2$ *induced by* σ_1.

The cell decomposition. The four train tracks τ_1, τ_2, τ_3, and τ_4 shown in the left-hand side of Figure 4 give a cell decomposition Y of $\mathrm{MF}(\mathrm{D}_3)$ as an integral cone complex. The top-dimensional cells are the $V_i = V(\tau_i)$, and the faces are given by the rays through four curves shown. For further explanation of this picture, see the book by Farb and the first author [11, Chapter 15] (the train tracks there are slightly different). Here we have another difference from our general setup, as the train tracks here are not Dehn–Thurston train tracks; on the other hand these train tracks are especially suited to the braid group. Also, this is a different structure on $\mathrm{MF}(\mathrm{D}_3) \cong \mathrm{MF}(S_{0,4})$ given in the introduction; the calculation can be carried out just as well with that setup.

The subdivisions and the matrices. The eight train tracks $\tau_{1,1}, \tau_{1,2}, \ldots, \tau_{4,1}, \tau_{4,2}$ shown in the right-hand side of Figure 4 give the common refinement of the collection of subdivisions Y_Γ of Y associated to Γ. Indeed, for each $f \in \Gamma^{\pm}$ and each τ_{ij} there is a k so that $f(V_{i,j}) \subseteq V_k$ and $f|V_{i,j}$ is linear. We illustrate this claim in one case in Figure 5 (for an example of a similar computation, see again the book by Farb and the first author). The figure shows $\tau_{1,2}$, $\sigma_1(\tau_{1,2})$, and τ_2, along with the weights corresponding to the xy-coordinates on $V_{1,2}$. We see from the figure that

$$\sigma_1(V_{1,2}) \subseteq V_2.$$

For each generator (or inverse of a generator) h, the subdivision Y_h lies between Y and the common subdivision. For instance, the subdivision Y_{σ_1} has top-dimensional cells $V_{1,1}$, $V_{1,2}$, V_2, V_3, and V_4. In each of the other three cases, the subdivision Y_h is again obtained by subdividing one cell of Y.

The calculations that allow us to check the inclusions $f(V_{i,j}) \subseteq V_k$ also allow us to compute the collection of matrices $\mathcal{A}$. From the weights in Figure 5 we see that the matrix $A_{1,2}^{\sigma_1}$ describing the action

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

of σ_1 on $V_{1,2}$ is

$$A_{1,2}^{\sigma_1} = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$$

Since the coordinates on $V_{1,2}$ are compatible with the coordinates on V_2 (that is, the inclusion is given by the identity matrix), the matrix $A_{1,2}^{\sigma_1}$ does (as required) describe the map $\sigma_1|V_{1,2}$ in terms of the V_1- and V_2-coordinates. As such, we can describe the action of σ_1 on its linear piece $V_{1,2}$ by the triple $(A_{1,2}^{\sigma_1}, V_1, V_2)$. The computations for the other cases are similar.

From the refinement given in Figure 4, we also obtain notation for each subdivision $Y_{\sigma_i^{\pm 1}}$. For example, Y_{σ_1} is the subdivision given by $V_{1,1}$, $V_{1,2}$, V_2, V_3, and V_4. While we will not list the subdivisions for the other seven generators and inverses of generators, the subdivisions are implicit from the calculations below: if we write $A_{1,2}^{\sigma_1}$ this signifies that $V_{1,2}$ is a linear piece for σ_1, and if we write $A_2^{\sigma_1}$ this signifies that V_2 is a linear piece for σ_1.

2.2. Applying the algorithm

As above, let $f = \sigma_2\sigma_1^{-1}\sigma_2\sigma_1\sigma_1\sigma_1$. Let c be the curve described by the vector $(1, 2)$ in V_1-coordinates (again, contrary to our general setup, this curve does not lie on a vertex ray). We first use the basic computation algorithm to give a detailed description of how to compute $f(c)$, and then use the modified version of the main algorithm to find the stretch factor and foliations.

Applying the basic computation algorithm. Thinking of f as a word w in Γ, we have $f_1 = \sigma_1$, $f_2 = \sigma_1$, etc. As in the description of the basic computation, we initialize by setting $c_0 = c$.

The following table gives the details of the steps of the basic computation for finding $f(c)$. In each row we give:

1. the step number ℓ,
2. the curve $c_{\ell-1}$ from the previous step,
3. the cell $V_{i,k}$ of the subdivision Y_Γ containing $c_{\ell-1}$,
4. the next letter f_ℓ in the word w,
5. the codomain V_j for $f_\ell|V_{i,k}$, and
6. the matrix $A_{i,k}^\ell$ so that $f_\ell|V_{i,k}$ is described by $(A_{i,k}^\ell, V_i, V_j)$.

Here is the table:

Table 1.

ℓ	$c_{\ell-1}$	$V_{i,k}$	f_ℓ	V_j	$A_{i,k}^\ell$	c_ℓ
1	$\binom{1}{2}$	$V_{1,2}$	σ_1	V_2	$\left(\begin{smallmatrix} -1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$	$\binom{1}{1}$
2	$\binom{1}{1}$	V_2	σ_1	V_3	$\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$	$\binom{1}{1}$
3	$\binom{1}{1}$	V_3	σ_1	V_4	$\left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$	$\binom{1}{2}$
4	$\binom{1}{2}$	V_4	σ_2	V_1	$\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$	$\binom{2}{1}$
5	$\binom{2}{1}$	V_1	σ_1^{-1}	V_1	$\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$	$\binom{3}{1}$
6	$\binom{3}{1}$	V_1	σ_2	V_2	$\left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$	$\binom{1}{4}$

(For $\ell > 1$ we use V_i instead of $V_{i,k}$ since the entire V_i is a linear piece for the corresponding f_ℓ.) From the table, we see that

$$f(c) = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \in V_2.$$

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

An aside. We can further see from the above computation that c lies in a linear piece for f described by a triple (A, V_1, V_2), where A is given by the product of the above matrices:

$$\begin{aligned} A &= A_1^{\sigma_2} A_1^{\sigma_1^{-1}} A_4^{\sigma_2} A_3^{\sigma_1} A_2^{\sigma_1} A_{1,2}^{\sigma_1} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 1 \\ -2 & 3 \end{pmatrix} \end{aligned}$$

We can check this by observing that

$$\begin{pmatrix} -1 & 1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

We will not use this description of the action of f on the linear piece containing c; we include it just to illustrate how the computations work.

Iterative-guess-and-check algorithm. We now describe in more detail the guess-and-check algorithm described at the start of the section. This method is the underpinning of some of the computer programs that predate our work. The idea of the algorithm is that, starting with $k = 1$, we "guess" that $f^k(c)$ lies in a Nielsen–Thurston eigenregion for f, and then

1. compute a triple (A, V_i, V_j) for the action of f on a linear piece containing the curve $f^k(c)$,
2. find pairs (λ, v) where $\lambda > 1$, where v lies in $V_i \cap V_j$, and where $f(v) = \lambda v$ (if $V_i = V_j$ such v can be computed as classical eigenvectors), and
3. check if one pair (λ, v) with $\lambda > 1$ and $v \geq 0$ is a PL-eigenvalue/eigenvector pair for f.

Since f has exactly one PL-eigenvalue/eigenvector pair with $\lambda > 1$ (up to scale), the last step implies that λ is the stretch factor and v represents the unstable foliation for f. If either the second or third step fails (meaning no pairs are found in the second step, or it is found in the third step that they are not PL-eigenvectors), we return to the first step, replacing $f^k(c)$ with $f^{k+1}(c)$. If f is pseudo-Anosov, then this process eventually terminates, since $f^k(c)$ must lie in a Nielsen–Thurston eigenregion for k large.

Again, the complexity of the iterative-guess-and-check algorithm is impossible to determine without an upper bound on the power of f required. Our Theorem 1.1 exactly gives such an upper bound, namely Q.

Applying the iterative-guess-and-check algorithm. We now apply the above algorithm to our specific $f \in \mathrm{B}_3$. As we will see, the algorithm terminates for $k = 1$. So this process terminates faster than if we were to apply the main algorithm, which would require computing $f^Q(c)$ with $Q = 2,561$.

(1) For the first step we apply the basic computation algorithm to the action of f on $f(c)$ and find the resulting acting matrix:

$$\begin{aligned} A &= A_2^{\sigma_2} A_3^{\sigma_1^{-1}} A_4^{\sigma_2} A_4^{\sigma_1} A_3^{\sigma_1} A_2^{\sigma_1} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \end{aligned}$$

This computation is done in the same way as the computation of $f(c)$. We do not include the corresponding table, but this can be inferred from the given product. During the computation, the curve $f(c) \in V_2$ gets mapped to V_3, to V_4, to V_4, to V_1, again to V_1, and back to V_2, in that order. In particular

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

$f(f(c))$ lies in V_2. This means that the result of the first step is the triple:

$$\left(\begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}, V_2, V_2\right)$$

(2) We now proceed to the second step. Since we have $V_i = V_j$, and we are guessing that $f(f(c))$ lies in a Nielsen–Thurston eigenregion for f, we compute the eigenvalue $\lambda > 1$ and corresponding eigenvector for the matrix A found in the first step. We find

$$\lambda = \frac{5+\sqrt{21}}{2} \quad \text{and} \quad v = \begin{pmatrix} 1+\sqrt{21} \\ 2 \end{pmatrix}.$$

(3) For the third and final step, we again use the basic computation algorithm to compute the action of f on a linear piece containing the eigenvector found in the last step. This time the computation turns out to be exactly the same as the computation in the first step, as the curves computed during the computation lie in exactly the same linear pieces for the six letters in the word for f. It follows that the pair (λ, v) found above is a PL-eigenvalue/eigenvector pair for f. As above, this means that λ is the stretch factor for f and the measured train track (τ_2, v) represents the unstable foliation.

In order to find coordinates for the stable foliation for f, we can repeat the whole process, except with f replaced by f^{-1}. We find that the stable foliation is represented by $(-3+\sqrt{21}, 6)$ in V_1-coordinates.

3. Slope

The main goal of this section is to prove Proposition 3.1, which bounds the diameter of the horizontality of a curve or leaf of a foliation. We also introduce other tools related to slope and horizontality. At the end of the section we give our application of slope to translation lengths in the curve graph, Corollary 3.8.

For the statement of Proposition 3.1, recall from the introduction that $Q_{\text{curve}}(S_{g,p}) = 2K^2$ and $K = K(S_{g,p}) = 4(3g-3_p)$. Also recall that $\mathcal{H}_f$ is the horizontality function coming from any singular Euclidean surface associated to a pseudo-Anosov f.

Proposition 3.1. *Let $S = S_{g,p}$ and let $f\colon S \to S$ be a pseudo-Anosov homeomorphism with stretch factor λ and a singular Euclidean structure X. If Z is a set of pairwise noncrossing saddle connections in $\bar{X}$, then*

$$\operatorname{diam} \mathcal{H}_f(Z) \leq Q_{\text{curve}}(S).$$

In our applications of Proposition 3.1 we will take Z to be the set of saddle connections appearing in the (flat) geodesic representative c_X of a curve c or a leaf of a foliation in S. This makes sense because the set of saddle connections appearing in the geodesic representative c_X of a curve or a leaf are pairwise noncrossing. More generally, Minsky and the third-named author of this paper proved that if c and d are two disjoint curves/leaves, then the geodesic representatives c_X and d_X consist of pairwise noncrossing saddle connections; see [22, Lemma 4.3].

Outline of the section. We begin in Section 3.1 with some preliminary definitions: completed universal covers, rectangles, and rectangle flexibility. The latter, denoted $\operatorname{flex}(X)$, describes the amount that rectangles in a singular Euclidean surface can be extended in the horizontal and vertical directions.

Next, in Section 3.2 we prove several lemmas that compare slopes of saddle connections in a fixed singular Euclidean surface. We first prove Lemma 3.2, a local slope disparity lemma, which bounds in terms of $\operatorname{flex}(X)$ the disparity between the slopes of saddle connections in a single rectangle (we obtain Lemma 3.2 as a consequence of Lemma 3.3, which compares the slope of a saddle connection to the slope of a rectangle containing it). We then promote Lemma 3.2 to Lemma 3.4, a global slope disparity lemma, which bounds in terms of $\operatorname{flex}(X)$ the disparity between the slopes of any noncrossing saddle connections in X.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

In Section 3.3 we consider singular Euclidean surfaces together with associated pseudo-Anosov maps. In Lemma 3.6 we relate $\mathrm{flex}(X)$ to the stretch factor of a pseudo-Anosov map. We then turn to the proof of Proposition 3.1. Since the slope of a curve is the set of slopes of a collection of noncrossing saddle connections, Proposition 3.1 will follow from Lemmas 3.4 and 3.6.

Finally, in Section 3.4, we apply the notion of horizontality to prove Corollary 3.8. This corollary bounds from below the stable translation length of the action of a pseudo-Anosov mapping class on the corresponding curve graph, extending the work of Gadre–Tsai.

3.1. Preliminaries

In this section, we introduce the definitions of completed universal covers, rectangles, and rectangle flexibility, along with several related notions.

Completed universal covers. Let X be a singular Euclidean surface and let $\bar{X}$ be the associated closed singular Euclidean surface, obtained from X by filling in the punctures. As in the work of Minsky and the third-named author, we denote by $\widehat{X}$ the metric completion of the universal cover of X. We refer to $\widehat{X}$ as the completed universal cover of X. There is an induced projection $\widehat{X} \to \bar{X}$. This map can be regarded as a branched cover, with infinite-fold ramification over the points of $\bar{X} \setminus X$.

The universal cover $\widetilde{X}$ can be identified with $\mathbb{H}^2$, and as such the completed universal cover $\widehat{X}$ can be identified with a subset (but not a subspace!) of the visual compactification of $\mathbb{H}^2$. The space $\widetilde{X}$ is a CAT(0) space. In particular, given a pair of points in $\widetilde{X}$, there is a unique geodesic connecting them. This geodesic may pass through arbitrarily many points of $\widehat{X} \setminus \widetilde{X}$.

Rectangles and saddle connections. Let X be a singular Euclidean surface with horizontal and vertical foliations $\mathcal{F}_h$ and $\mathcal{F}_v$. A rectangle in X or $\bar{X}$ is defined to be an isometric immersion $R\colon E \to \bar{X}$, where

1. E is a Euclidean rectangle,
2. each horizontal/vertical segment in E maps to a horizontal/vertical segment, and
3. $R^{-1}(\Sigma) \subseteq \partial E$ (recall Σ is the set of singular points).

While R is an immersion, it is always possible to lift R to $\widehat{X}$ and any such lift will be an embedding. In all of our figures, the rectangles will be shown as embedded; we can (safely) imagine that these pictures describe lifts to $\widehat{X}$.

By a diagonal of R we mean the R-image of a diagonal of E. When we say that a point $x \in X$ lies in the boundary or interior of R, we mean that $R^{-1}(x)$ lies in the boundary or interior of E, respectively. The height and width of R are defined to be the height and width of E.

We say that a saddle connection σ in $\bar{X}$ is rectangle spanning if it is the diagonal of a rectangle. In this case the rectangle is denoted R_σ and is called the spanning rectangle for σ. The height and width of a rectangle-spanning saddle connection are defined to be the height and width of its spanning rectangle.

Minsky and the third-named author proved that if X is a singular Euclidean surface associated to a pseudo-Anosov mapping class f, the triangulations of X by rectangle-spanning saddle connections are exactly the sections of Agol's veering triangulation [22, Section 3].

Slopes of rectangles and flexibility. Let X be a singular Euclidean surface. As in the introduction, we define the slope of a rectangle $R \subseteq X$ as $\mathrm{slope}(\delta)$, where δ is either diagonal of R (the diagonals have equal slope). Also, we say that a rectangle $\hat{R}$ is an extension of R if it is obtained by extending R in one direction (up, down, left, or right). Finally, we define the rectangle flexibility of X as

$$\mathrm{flex}(X) = \sup_{(\sigma, \hat{R}_\sigma)} \max\left\{ \frac{\mathrm{slope}(\sigma)}{\mathrm{slope}(\hat{R}_\sigma)}, \frac{\mathrm{slope}(\hat{R}_\sigma)}{\mathrm{slope}(\sigma)} \right\},$$

where σ is a rectangle-spanning saddle connection in X and $\hat{R}_\sigma$ is an extension of the spanning rectangle R_σ.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

Finite flexibility versus the existence of horizontal and vertical saddle connections. For a singular Euclidean surface X we have the following implication:

$$\operatorname{flex}(X) < \infty \Longrightarrow \bar{X} \text{ has no horizontal or vertical saddle connection}$$

Much of our work in this section relies on the work of Minsky and the third-named author of this paper [22]. In their work on singular Euclidean surfaces, they assume the surfaces do not have any horizontal or vertical saddle connections. As a proxy for this assumption, we will assume the condition that X has finite flexibility. It is a straightforward consequence of the definition of flexibility that this implies the nonexistence of vertical and horizontal saddle connections. The converse of the implication is also true, but we will not require it.

A singular Euclidean structure for a pseudo-Anosov homeomorphism has no vertical or horizontal saddle connections. Such a structure also has finite flexibility by Lemma 3.6 (or by the converse of the above implication). We will use this without mention below in order to apply our lemmas about singular Euclidean structures with finite flexibility to singular Euclidean structures associated to pseudo-Anosov homeomorphisms.

While Lemmas 3.2, 3.3, and 3.4 hold vacuously in the case where the flexibility is infinite, we state them in the case of finite flexibility for the sake of clarity.

3.2. Slope disparity lemmas

The main goals of this section are to prove Lemmas 3.2 and 3.4, the local and global slope disparity lemmas. In what follows, we will write $F = F(X)$ for the quantity $2\operatorname{flex}(X) - 1$; this quantity will appear often in what follows.

Local slope disparity. The next lemma bounds the disparity between slopes of saddle connections in a given rectangle.

Lemma 3.2. *Let X be a singular Euclidean surface with finite flexibility $F = F(X)$, let R be a rectangle in X, and let σ_0 and σ_1 be saddle connections contained in R. Then*

$$\operatorname{slope}(\sigma_0) \leq F^2 \operatorname{slope}(\sigma_1).$$

Lemma 3.2 is an immediate consequence of the following lemma, which compares the slope of a rectangle to the slopes of saddle connections contained within.

Lemma 3.3. *Let X be a singular Euclidean surface with finite flexibility $F = F(X)$, and let R be a rectangle in X and σ a saddle connection in R. Then*

$$\frac{1}{F} \leq \frac{\operatorname{slope}(\sigma)}{\operatorname{slope}(R)} \leq F.$$

Proof. We prove the second inequality. The first is proved with a symmetric argument. Let R_σ^h be the maximal rectangle obtained by extending the spanning rectangle R_σ horizontally in both directions. Note that σ connects the horizontal edges of R_σ^h.

As $\operatorname{slope}(R_\sigma^h) \leq \operatorname{slope}(R)$ and $F = 2\operatorname{flex}(X) - 1$, it suffices to prove that

$$\frac{\operatorname{slope}(\sigma)}{\operatorname{slope}(R_\sigma^h)} \leq 2\operatorname{flex}(X) - 1.$$

Since ratios of slopes are invariant under horizontal and vertical scaling, we may assume for simplicity that

$$\operatorname{height}(R_\sigma) = \operatorname{width}(R_\sigma) = 1.$$

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

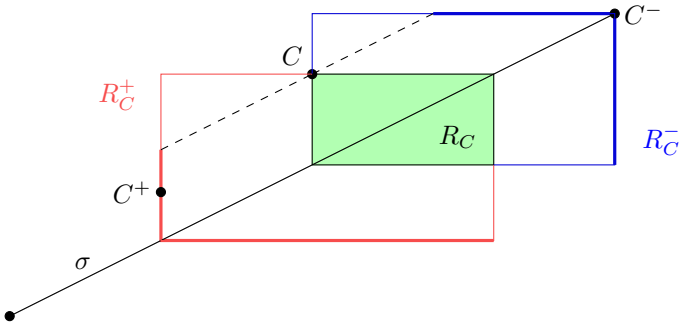


Figure 6. *The rectangles R_C, R_C^+, and R_C^- from the proof of Lemma 3.5.*

Let R_σ^- and R_σ^+ be the maximal extensions of R_σ to the left and right. By definition, we have

$$\operatorname{width}(R_\sigma^\pm) \leq \operatorname{flex}(X).$$

If E, E^h, E^-, and E^+ are the domains of R_σ, R_σ^h, R_σ^-, and R_σ^+, respectively, then $E^- \cup E^+ = E^h$ and $E^- \cap E^+ = E$. Therefore

$$\begin{aligned}\frac{\operatorname{slope}(\sigma)}{\operatorname{slope}(R_\sigma^h)} &= \frac{1}{\operatorname{slope}(R_\sigma^h)} = \operatorname{width}(R_\sigma^h) \\ &= \operatorname{width}(R_\sigma^+) + \operatorname{width}(R_\sigma^-) - \operatorname{width}(R_\sigma) \leq 2\operatorname{flex}(X) - 1,\end{aligned}$$

as desired. □

Global slope disparity. Our next goal is to prove the following lemma, which bounds the disparity between slopes of noncrossing saddle connections. In the statement, the constant $K = K(S_{g,p})$ is the number $4(3g - 3 + p)$ from the introduction, namely, our upper bound for the number of separatrices in a foliation on $S_{g,p}$.

Lemma 3.4. *Let X be a singular Euclidean surface with finite flexibility $F = F(X)$. If σ_0 and σ_1 are noncrossing saddle connections in X, then*

$$\operatorname{slope}(\sigma_0) \leq F^{2K} \operatorname{slope}(\sigma_1).$$

Our proof of Lemma 3.4 requires a subordinate lemma, Lemma 3.5 below. Before stating that lemma, we recall some preliminaries from the work of Minsky and the third-named author, in particular the notion of a rectangle hull; see [22, Section 4.1]. We also introduce the notion of a rectangle associated to a point.

Let X be a singular Euclidean surface with finite flexibility. For any saddle connection σ of X, let $\mathcal{R}(\sigma)$ be the collection of rectangles in X that are maximal with respect to the property that their diagonals lie along σ. We refer to $\mathcal{R}(\sigma)$ as the rectangle casing for σ.

The rectangle hull $\mathbf{r}(\sigma)$ is defined to be the collection of saddle connections contained in some rectangle of $\mathcal{R}(\sigma)$. For example, the rectangle hull of the saddle connection in Figure 6 contains the four saddle connections that join the singularities in the boundary of R^+ or the boundary of R^-. Every rectangle hull is connected (the rectangles in $\mathcal{R}(\sigma)$ are ordered along σ and consecutive rectangles share a singularity). If σ_0 and σ_1 are noncrossing saddle connections, then the saddle connections of $\mathbf{r}(\sigma_0) \cup \mathbf{r}(\sigma_1)$ are also noncrossing [22, Lemma 4.1].

If C is a singular point in a rectangle $R \in \mathcal{R}(\sigma)$ and does not lie on σ, then we define the associated rectangle R_C as the rectangle in R with corner C and diagonal along σ. In Figure 6 we show a representative configuration.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

Let R_C^- and R_C^+ be the rectangles of $\mathcal{R}(\sigma)$ obtained by sliding one of the points of $R_C \cap \sigma$ along σ. The rectangle R_C is contained in both of these. Also, it is possible that one or both is equal to R_C.

The set $\mathcal{R}(\sigma)$ is finite, because the diagonals of the elements of $\mathcal{R}(\sigma)$ form a pairwise nonnested cover of σ by closed intervals. Also, each element of $\mathcal{R}(\sigma)$ contains at most one singularity on each edge, since we have assumed that X has no horizontal or vertical saddle connections. In particular, there are finitely many R_C as above.

Lemma 3.5. *Let X be a singular Euclidean surface with finite flexibility. For any saddle connection σ, there are rectangle-spanning saddle connections σ^+, σ^- in the rectangle hull $\mathbf{r}(\sigma)$ such that*

$$\operatorname{slope}(\sigma^-) \leq \operatorname{slope}(\sigma) \leq \operatorname{slope}(\sigma^+).$$

Proof. As above, there are finitely many rectangles R_C associated to singularities in the rectangles of $\mathcal{R}(\sigma)$. Thus there is such an R_C of largest area.

Without loss of generality, we assume that σ passes through the bottom-left and top-right corners of R_C and that C lies above (or at an endpoint of) σ (if needed, we precompose R with a symmetry of its domain rectangle E). Let R_C^+ and R_C^- denote the elements of $\mathcal{R}(\sigma)$ that contain R_C and are obtained by sliding, respectively, the bottom-left and top-right corners of R_C along σ. As above, it is possible for one or both of these two to equal R_C. See Figure 6 for an illustration.

There must be a singularity C^+ on the left or bottom boundary of R_C^+ and also a singularity C^- on the top or right boundary of R_C^-, for otherwise, R_C^+ and R_C^- would not be maximal (as demanded by the definition of the rectangle hull).

Let ℓ be the line segment in $R_C^+ \cup R_C^-$ that passes through C and is parallel to σ. The singularities C^+ and C^- lie on the same side of ℓ as σ. Indeed, if not, there would be a translate of $R_C^\pm$ (in the direction of σ) that strictly contains R_C, violating the maximality of the latter.

Let σ^+ and σ^- be the saddle connections connecting C^+ and C^- to C within R_C^+ and R_C^-. We have by definition that σ^+ and σ^- lie in $\mathbf{r}(\sigma)$. By the previous paragraph they satisfy $\operatorname{slope}(\sigma^-) \leq \operatorname{slope}(\sigma) \leq \operatorname{slope}(\sigma^+)$, as desired. □

We now turn toward the proof of Lemma 3.4. In the proof, we use the fact that $K(S)$, our upper bound for the number of separatrices in a foliation on S, is also an upper bound for the number of triangles in a triangulation of S by saddle connections. Indeed, the map associating each separatrix to the triangle it initially enters is bijective.

Proof of Lemma 3.4. We proceed in two steps, first proving the lemma in the case of rectangle-spanning saddle connections and then in the general case.

To this end, first suppose that σ_0 and σ_1 are noncrossing rectangle-spanning saddle connections in X. Minsky and the third-named author proved that any collection of noncrossing, rectangle-spanning saddle connections can be extended to a triangulation by rectangle-spanning saddle connections [22, Lemma 3.2]. Consider then such a triangulation of X. As above, the number of triangles is bounded above by $K = K(S)$. Thus, any sequence of distinct triangles (T_i) has length at most K. It further follows that there is a sequence of edges of the triangulation $e_0 = \sigma_0, e_2, \ldots, e_n = \sigma_1$ with $n \leq K$ and with each consecutive pair contained in a triangle. Since each triangle is contained in a rectangle, the special case now follows from Lemma 3.2.

We now proceed to the general case. Let σ_0 and σ_1 be arbitrary noncrossing saddle connections. For $i = 0, 1$ let $\sigma_i^\pm$ be the rectangle-spanning saddle connections given by Lemma 3.5. Minsky and Taylor proved that if two saddle connections are noncrossing, then the elements of their rectangle hulls are pairwise noncrossing [22, Lemma 4.1]. Therefore, by Lemma 3.5 and the special case of this lemma, we have

$$\frac{\operatorname{slope}(\sigma_0)}{\operatorname{slope}(\sigma_1)} \leq \frac{\operatorname{slope}(\sigma_0^+)}{\operatorname{slope}(\sigma_1^-)} \leq F^{2K},$$

as desired. □

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

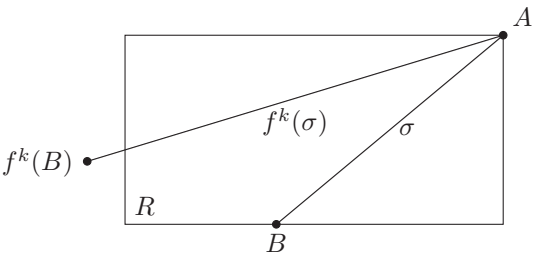


Figure 7. *The saddle connections σ and $f^k(\sigma)$ from the proof of Lemma 3.6.*

3.3. Slopes and stretch factors

In this section, we incorporate the slope disparity lemmas together with basic properties of pseudo-Anosov maps in order to prove Proposition 3.1. We begin with a comparison of rectangle flexibilities and stretch factors.

Flexibility versus stretch factor. The following lemma bounds the rectangle flexibility of a singular Euclidean surface corresponding to a pseudo-Anosov map in terms of the stretch factor. Again, $K = K(S_{g,p}) = 4(3g - 3 + p)$.

Lemma 3.6. *Let $S = S_{g,p}$ and let $f\colon S \to S$ be pseudo-Anosov with stretch factor λ and singular Euclidean structure X. Then*

$$\operatorname{flex}(X) \leq \lambda^K.$$

Proof. Let $\hat{R}_\sigma$ be an extension of a spanning rectangle for a saddle connection σ in X. And let us denote $\hat{R}_\sigma$ by R. By symmetry, we may assume that R is a horizontal extension, as in Figure 7 (for the other case we replace f with f^{-1}). Say that the endpoints of σ are the singularities A and B. It must be the case that one of these, say A, is at a corner of R. The singularity B must then lie on a horizontal edge of R. We need to show two inequalities:

$$\frac{\operatorname{slope}(R)}{\operatorname{slope}(\sigma)} \leq \lambda^K \quad \text{and} \quad \frac{\operatorname{slope}(\sigma)}{\operatorname{slope}(R)} \leq \lambda^K.$$

The first inequality is immediate since $\operatorname{slope}(R) \leq \operatorname{slope}(\sigma)$ and $\lambda > 1$. We proceed to the second inequality. Choose $k \leq K$ so that f^k fixes A as well as the separatrices of the horizontal and vertical foliations at A. Since R contains no singularities, $f^k(B)$ cannot lie in the interior of R. Thus

$$\operatorname{width}(R) \leq \operatorname{width}(f^k(\sigma)).$$

On the other hand, we have

$$\operatorname{height}(R) = \operatorname{height}(\sigma) = \lambda^k \operatorname{height}(f^k(\sigma)).$$

Dividing the first and last expressions here by the expressions in the previous inequality, we obtain

$$\operatorname{slope}(R) \geq \lambda^k \operatorname{slope}(f^k(\sigma)).$$

Since $\operatorname{slope}(\sigma) = \lambda^{2k} \operatorname{slope}(f^k(\sigma))$, we have

$$\frac{\operatorname{slope}(\sigma)}{\operatorname{slope}(R)} \leq \frac{\lambda^{2k} \operatorname{slope}(f^k(\sigma))}{\lambda^k \operatorname{slope}(f^k(\sigma))} = \lambda^k \leq \lambda^K,$$

as desired. □

Penner's lower bound. Penner proved the following lemma [26].

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

Lemma 3.7 (Penner). *If $f \in \mathrm{Mod}(S)$, then*

$$2 \leq \lambda^K$$

where λ is the stretch factor of f and $K = K(S)$.

Penner stated the equivalent formulation $\log \lambda(f) \geq (\log 2)/K$. In the proof of Proposition 3.1 we will use the inequality $4^K \leq \lambda^{2K^2}$, obtained from Penner's inequality by squaring both sides and raising to the Kth power.

Completing the proof. We are now ready for the proof of Proposition 3.1, which gives an upper bound on the diameter of a set of noncrossing saddle connections.

Proof of Proposition 3.1. As in the statement of the proposition, S is a surface, $f\colon S \to S$ is pseudo-Anosov with stretch factor λ, X is an associated singular Euclidean surface, and Z is a collection of noncrossing saddle connections. Also recall from the introduction that $Q_{\mathrm{curve}} = Q_{\mathrm{curve}}(S) = 2K^2$. Let $F = F(X)$ be the flexibility. The conclusion of the proposition is equivalent to the inequality

$$\frac{\max \mathrm{slope}(Z)}{\min \mathrm{slope}(Z)} \leq (\lambda^2)^{Q_{\mathrm{curve}}}.$$

We proceed to the proof of this inequality. Let σ_0 and σ_1 be any two saddle connections of Z. We have

$$\begin{aligned}
\mathrm{slope}(\sigma_0) &\leq \mathrm{slope}(\sigma_1)F^{2K} && \text{(Lemma 3.4)}\\
&= \mathrm{slope}(\sigma_1)(2\,\mathrm{flex}(X) - 1)^{2K} && \text{(definition of } F\text{)}\\
&\leq \mathrm{slope}(\sigma_1)(2\lambda^K - 1)^{2K} && \text{(Lemma 3.6)}\\
&< \mathrm{slope}(\sigma_1)4^K\lambda^{2K^2} && (2\lambda^K - 1 < 2\lambda^K)\\
&\leq \mathrm{slope}(\sigma_1)(\lambda^2)^{2K^2} && \text{(Lemma 3.7)}
\end{aligned}$$

as required. □

3.4. Application to stable translation lengths in the curve graph

As above, Proposition 3.1 gives an upper bound of Q_{curve} on the diameter of the horizontality of a collection of pairwise noncrossing saddle connections. This proposition has the following corollary. In the statement, $\mathcal{C}(S)$ is the curve graph for S and $\tau_{\mathcal{C}(S)}(f)$ is the stable translation length for the action of a pseudo-Anosov f on $\mathcal{C}(S)$. The second statement of the corollary is a (slight) quantitative strengthening of a special case of a result of Gadre–Tsai [15, Theorem 1.1].

Corollary 3.8. *Let $S = S_{g,p}$ and let $f : S \to S$ be pseudo-Anosov. The horizontality function defines a Q_{curve}-coarse Lipschitz map*

$$\mathcal{H}_f : \mathcal{C}(S) \to \mathbb{R}.$$

Hence, for any pseudo-Anosov $f \in \mathrm{Mod}(S)$, we have

$$\tau_{\mathcal{C}(S)}(f) \geq \frac{1}{Q_{\mathrm{curve}}}.$$

For closed surfaces, Gadre–Tsai proved that for $g \geq 2$ and any pseudo-Anosov $f \in \mathrm{Mod}(S_g)$ we have

$$\tau_{\mathcal{C}(S_g)}(f) \geq \frac{1}{162(2g-2)^2 + 30(2g-2)}.$$

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

Since $Q_{\text{curve}} = 2K^2 < 162(2g-2)^2 + 30(2g-2)$ for $g > 1$, our Corollary 3.8 is a strengthening of this. For surfaces with punctures, Gadre–Tsai give a lower bound that is stronger than the one given by our corollary.

4. Approximating rectangles

The goal of this section is to prove Lemma 4.1. This is the most technical result in the paper. Let X be a singular Euclidean surface and let $\widehat{X}$ be its completed universal cover, as in Section 3. What Lemma 4.1 tells us is that if γ_1 and γ_2 are geodesics in $\widehat{X}$ that cross each other and also cross a singularity-free horizontal leaf ℓ, then there is a maximal rectangle R in $\widehat{X}$ with the following properties:

- γ_1 and γ_2 cross the horizontal sides of R, and
- any saddle connection in X with sufficiently small slope has a path lift that crosses the vertical sides of R.

Before stating Lemma 4.1, we introduce the notion of a k-approximating rectangle.

Approximating rectangles. As above let X be a singular Euclidean surface with $\operatorname{flex}(X) < \infty$ and let $\widehat{X}$ be its completed universal cover. Let γ_1 and γ_2 be intersecting geodesics in $\widehat{X}$ whose endpoints (if any) are singular points of $\widehat{X}$ (all points of $\widehat{X} \setminus \overline{X}$ are singular).

Let ℓ be a bi-infinite horizontal leaf in $\widehat{X}$ that intersects both γ_1 and γ_2. We say a rectangle R in $\widehat{X}$ is a *k-approximating rectangle* for the triple $(\gamma_1, \gamma_2, \ell)$ if

- R is maximal (i.e., it contains a singularity in each edge),
- ℓ contains an interior horizontal leaf of R,
- γ_1 and γ_2 connect the horizontal edges of R, and
- $k \operatorname{slope}(R) \geq \min\{\operatorname{slope}(\gamma_1) \cup \operatorname{slope}(\gamma_2)\}$.

As in Section 3, we write $F(X)$ for the quantity $2\operatorname{flex}(X) - 1$ when X is a singular Euclidean surface. Also, in what follows, by a lift of a saddle connection we will mean a path lift.

Lemma 4.1. *Let X be a singular Euclidean surface with $\operatorname{flex}(X) < \infty$. Let γ_1 and γ_2 be intersecting geodesics in $\widehat{X}$ whose endpoints (if any) are singular points of $\widehat{X}$. Let ℓ be a singularity-free horizontal leaf in $\widehat{X}$ that intersects both γ_1 and γ_2. Let $F = 2\operatorname{flex}(X) - 1$ and let $K = K(X)$ be the separatrix complexity. Then the following hold.*

1. *There exists an F^5-approximating rectangle R for $(\gamma_1, \gamma_2, \ell)$.*
2. *If R is an F^5-approximating rectangle for $(\gamma_1, \gamma_2, \ell)$ and Z is a collection of pairwise noncrossing saddle connections in X with*

$$\min \operatorname{slope}(Z) < F^{-(2K+6)} \operatorname{slope}(\gamma_i) \text{ for } i \in \{1, 2\}$$

then there is a $\sigma \in Z$ and a lift $\tilde{\sigma}$ that crosses the vertical edges of R.

4.1. The existence of approximating rectangles

For the proof of the first statement of Lemma 4.1 – on the existence of approximating rectangles – we introduce three ingredients: a basic fact about triangles in singular Euclidean surfaces, developing maps, and a graph of rectangles and rectangle-spanning saddle connections.

Triangles in singular Euclidean surfaces. In the proof of Lemma 4.1, we will use the fact that a triangle in a (simply connected) singular Euclidean surface – a region bounded by three geodesics – has no singularities in its interior. This follows readily from the Gauss–Bonnet formula for singular Euclidean surfaces with geodesic boundary, namely:

$$2\pi\chi(Y) = -\pi\left(\sum_{p_k} \deg(p_k)\right) + \sum_i (\pi - \theta_i),$$

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

where the θ_i are the interior angles on the boundary and each $\deg(p_k) \geq 1$ is the degree of the interior singularity p_k (the degree is two less than the number of prongs); see [28, Theorem 3.3].

Developing maps. Let X be a singular Euclidean surface. By a disk in $\widehat{X}$ we mean an isometric embedding of a closed polygonal disk in the Euclidean plane. Examples of disks are (the images of) rectangles. It follows from the Gauss–Bonnet formula that a triangle (as in the previous paragraph) is a disk in this sense.

Suppose we have a collection of disks $D_0, D_1, \ldots$ in $\widehat{X}$. Suppose that each D_i with $i > 0$ has the property that its intersection with $D_0 \cup \cdots \cup D_{i-1}$ is connected and contains a nonempty open set. In this case there is a developing map from $\cup D_i$ to the Euclidean plane that respects horizontal and vertical directions and measures. This map is defined inductively. First we make a choice of embedding of D_0 in the plane; this choice is unique up to translation, rotation by π, and reflection through a horizontal or vertical line. Then each subsequent D_i is mapped to the plane in a way that agrees with the intersection with $D_0 \cup \cdots \cup D_{i-1}$. Given a choice of embedding of D_0, the rest of the developing map is determined.

In the proof of Lemma 4.1, we will use the (almost) uniqueness of the developing map as a way of relating certain configurations of triangles and rectangles in a singular-Euclidean surface to a configuration of disks in the Euclidean plane. Once we fix a developing map, it makes sense to use words like "above" and "to the left" when describing features of the configuration of disks.

The graph of rectangles and rectangle-spanning saddle connections. Let X be a singular Euclidean surface. We define $\mathcal{G}_X$ to be the graph whose vertices are the rectangles in X and the rectangle-spanning saddle connections in X. We connect a rectangle-spanning saddle connection σ to a rectangle R if σ is contained in R. It follows immediately from Lemma 3.3 that if $v_0, \ldots, v_n$ is a path in $\mathcal{G}_X$, then

$$F^n \operatorname{slope}_X(v_0) \geq \operatorname{slope}_X(v_n).$$

In the proof of Lemma 4.1 we will bound the slope of the approximating rectangle by constructing a path in the graph $\mathcal{G}_{\widehat{X}}$ associated to $\widehat{X}$. We will take v_0 to be the approximating rectangle R and v_n to be a saddle connection of γ_1 or γ_2 (or a saddle connection with slope greater than some such saddle connection).

Proof of Lemma 4.1(1). We break the proof into two main parts. In the first part we construct a rectangle R satisfying the first three conditions of the definition of an approximating rectangle, and in the second part we show that R satisfies the slope condition. We begin with some setup. Throughout, the reader may refer to Figure 8.

Setup. Denote the unique points of intersection $\ell \cap \gamma_1$ and $\ell \cap \gamma_2$ by x_1 and x_2 (these points are unique because two geodesics must intersect in a connected union of saddle connections and because ℓ contains no singularities). The geodesics γ_1, γ_2 and ℓ bound a triangle Δ. As above, the interior of Δ is free of singularities. We allow for the case where $x_1 = x_2$; in this case Δ is degenerate, namely, it is the single point x_1. In what follows, we assume that we are in the generic case where $x_1 \neq x_2$, and so Δ has nonempty interior.

As per Figure 8, we may choose a developing map for Δ so that Δ lies above ℓ. We may also assume that γ_1 and γ_2 lie to the left and right of Δ, respectively. These two choices determine the developing map up to translation.

Below, we will define rectangles $R_1, \ldots, R_5$, and at each stage it will be true that the union satisfies the criterion for the existence of a developing map, which is at each stage uniquely defined as an extension of the developing map defined on Δ. The images of the R_i and Δ under the developing map (at least in a generic case) are shown in Figure 8.

Part 1: Constructing R. In order to construct R, we first construct three intermediate rectangles, R_1, R_2, and R_3.

For each $i \in \{1, 2\}$, we define R_i to be a rectangle in $\mathcal{R}(\gamma_i)$ satisfying the property that R_i contains singularities on either side of ℓ. To explain why such rectangles exist, we fix some $i \in \{1, 2\}$. Since ℓ intersects γ_i, it crosses some saddle connection σ_i of γ_i. As ℓ contains no singular points, the two singularities of σ_i lie on opposite sides of γ_i. The rectangle hull $\mathbf{r}(\sigma_i)$ is a connected graph in $\widehat{X}$

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

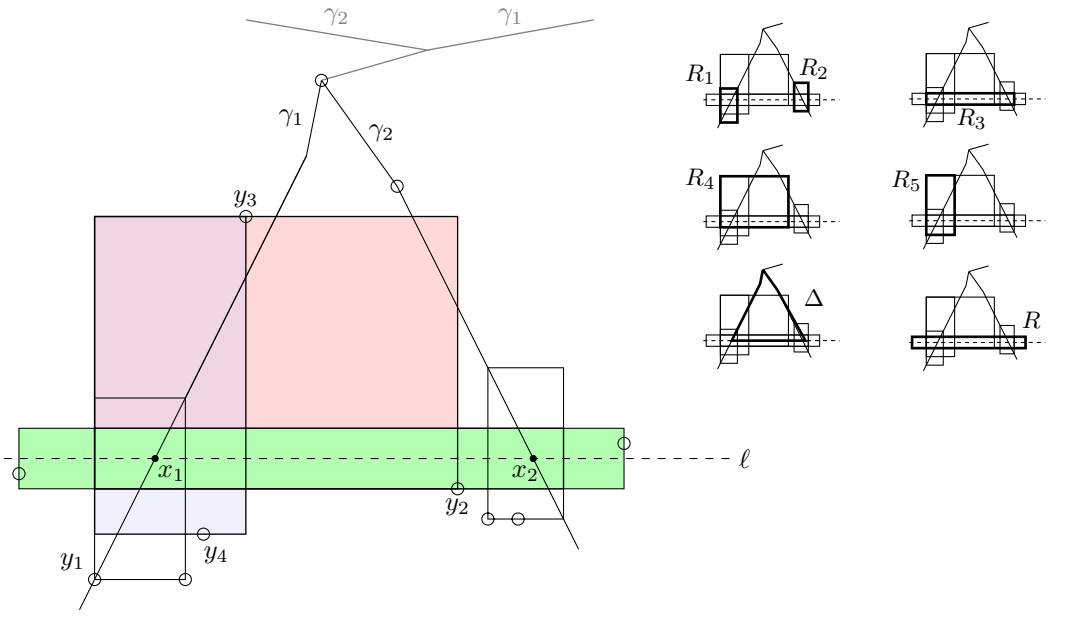


Figure 8. *The polygons in the proof of Lemma 4.1. The right column shows a key highlighting the individual rectangles and the triangle.*

containing the two singular points of σ_i. Therefore, there is a saddle connection ρ_i of $\mathbf{r}(\sigma_i)$ that connects singular points on opposite sides of ℓ. The rectangle $R_i \in \mathcal{R}(\sigma_i)$ associated to ρ_i satisfies the desired property.

We note that Figure 8 does not accurately reflect the case where $R_1 \cap R_2 \neq \emptyset$. But our arguments still apply in this case.

We now proceed to construct the rectangle R_3. Let ℓ_0 be the minimal segment of ℓ containing $R_1 \cap \ell$ and $R_2 \cap \ell$. Then let R_3 be the rectangle that is minimal with respect to the following properties:

- ℓ_0 is a horizontal leaf of R_3 and
- R_3 contains singularities on both its top and bottom edges.

The rectangle R_3 has a singularity on both of its horizontal edges (possibly at the corners), but not in the interiors of its vertical edges. Because R_1 and R_2 have singularities on both sides of ℓ, the endpoints of both vertical edges of R_3 are contained in $R_1 \cup R_2$, as in Figure 8.

Finally, let R be the rectangle that is maximal with respect to the property that the horizontal edges of R_3 are contained in R. In other words, R is the maximal horizontal extension of R_3.

To complete the first part of the proof, we check that R satisfies the first three conditions in the definition of an approximating rectangle.

- Since R_3 contains singularities on both of its horizontal edges and R is a maximal horizontal extension of R_3, we have that R is maximal.
- Since R_3 contains a segment of ℓ in its interior and R contains R_3, we have that R contains a segment of ℓ in its interior.
- Since R_1 and R_2 both contain singularities on both sides of ℓ, the vertical edges of R_3 are contained in the vertical edges of R_1 and R_2. Then, since the diagonals of R_1 and R_2 are contained in γ_1 and γ_2, respectively, it follows that γ_1 and γ_2 cross the horizontal edges of R_3, hence R, as desired.

This completes the first part of the proof.

Part 2: Estimating the slope. Our strategy for estimating the slope of R, and hence verifying the fourth condition of a k-approximating rectangle, is to find a path in $\mathcal{G}_{\widehat{X}}$ from R to a saddle connection that has slope greater than or equal to the slope of some saddle connection of γ_1 or γ_2. Since the goal is

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

to show that R is an F^5-approximating rectangle, it suffices (as above) to find such a path whose length is bounded above by 5. To this end, we start by defining singular points y_1, y_2, y_3, and y_4 and rectangles R_4 and R_5; each of the saddle connections we use in our path will be of the form $y_i y_j$.

- Let y_1 and y_2 be the singularities on the horizontal edges of R_3 that lie above and below ℓ, respectively.

These exist by the definition of R_3, and they are unique since $\widehat{X}$ has no horizontal saddle connections.

Since Δ, R_1, and R_2 have no singularities in their interiors, it follows that y_1 lies in the boundary of R_1 or R_2. Up to relabeling the γ_i, we may assume that it lies in the boundary of R_1, as in Figure 8. (Assuming $R_1 \neq R_2$ we have no further flexibility in the choice of developing map.) Because Δ has no singularities in its interior, y_1 must lie on the left of R_1 (as in Figure 8) or the top of R_1.

- Let R_4 be the rectangle that is maximal with respect to the properties that it contains y_2 at its bottom-right corner and y_1 along its left edge.
- Let y_3 be the singularity on the top edge of R_4.
- Let R_5 be the rectangle that is maximal with respect to the properties that it contains y_3 at its top-right corner and also contains y_1 along its left edge.
- Let y_4 be the singularity on the bottom edge of R_5.

In other words, R_4 is obtained from the rectangle spanned by the saddle connection y_1y_2 by extending it upward as far as possible and R_5 is obtained from the rectangle spanned by y_1y_3 by extending it downward as far as possible.

To complete the proof, we treat three cases. In each case we will choose a particular saddle connection σ of one of the γ_i. We will then find a saddle connection $y_i y_j$ or a rectangle R_i with slope greater than or equal to that of σ and with distance at most 5 from R in $\mathcal{G}_{\widehat{X}}$.

Case 1: y_3 lies to the left of γ_1. For this case let σ be the saddle connection of γ_1 that crosses ℓ. There are two slightly different subcases, according to whether y_1 lies on the left-hand edge or the top edge of R_1. We begin by treating the former case in detail and then explaining the minor changes needed for the latter case.

In both subcases, it follows from the definition of R_4 that y_1 and y_3 lie on the left-hand and top edges of R_4, respectively. It then follows that y_1 lies on the left-hand edge of R_5, and that y_3 is the top-right corner of R_5. It must be that y_3, and hence the top edge of R_5, lies outside Δ. In particular, it must be that σ intersects the right-hand edge of R_5. It also follows from the definitions that y_4 lies on the bottom edge of R_5.

For the case where y_1 lies on the left-hand edge of R_1, we claim that the following hold:

- σ intersects the right-hand edge of R_5
- y_3 is the top-right corner of R_5,
- y_4 lies on the bottom edge of R_5,
- σ intersects the bottom edge of R_5,
- along the bottom edge of R_5, y_4 lies to the right of σ.

We already verified the first three items. Since R_4 contains no singularities in its interior and since R_1 contains a singularity below ℓ, it must be that the bottom of R_5 passes through R_1, which gives the fourth item. As R_1 contains no singularities in the interior, and has no singularities on its left edge besides y_1, the fifth and final item follows.

It follows from the claim that

$$\operatorname{slope}(y_3y_4) \geq \min \operatorname{slope}(\gamma_1).$$

The path

$$R, y_1y_2, R_4, y_1y_3, R_5, y_3y_4$$

in $\mathcal{G}_{\widehat{X}}$ has length 5, and so the lemma follows in the case where y_1 lies on the left-hand edge of R_1.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

For the case where y_1 lies in the (interior of the) top edge of R_1, the proof is very similar. The changes are that the left-hand edge of R_5 does not overlap the left-hand edge of R_4, and that the bottom edge of R_5 can be lower than the bottom edge of R_1 (this can happen when the singularity of R_1 below ℓ lies to the left of R_5). The fourth and fifth items above become:

- σ intersects the bottom edge or the left-hand edge of R_5,
- if σ intersects the bottom edge of R_5, y_4 lies to the right of σ along this edge.

We can think of the case where σ crosses the left-hand edge of R_5 as being a case where σ crosses the bottom edge of R_5 very far to the left; then these two items are morally equivalent to their counterparts in the first subcase. As such, we still obtain that $\operatorname{slope}(y_3y_4) \geq \min \operatorname{slope}(\gamma_1)$, and the proof concludes as in the first subcase.

Case 2: y_3 lies to the right of γ_2, y_2 lies to the right of γ_2. For this case let σ be the saddle connection of γ_2 that crosses ℓ. We claim that in this case y_2y_3 is a saddle connection contained in R_2. Since R_2 contains no singularities in its interior, it follows from the definition of y_2 that y_2 must lie along the right-hand edge of R_2. It then follows from the maximality condition on R_2 and the fact that Δ contains no singularities in its interior that R_2 has a singularity along its top edge. From the definition of y_3, the fact that R_2 contains no singularities in its interior, and the assumption that y_3 lies to the right of γ_2, it must be that the singularity on the top edge of R_2 is nothing other than y_3. The claim follows.

Given the claim, we have the following path in $\mathcal{G}_{\widehat{X}}$:

$$R, y_1y_2, R_4, y_2y_3, R_2.$$

Since $\operatorname{slope}(R_2)$ is equal to the slope of the saddle connection of γ_2 comprising the diagonal of R_2, the lemma follows in this case.

Case 3: y_3 lies to the right of γ_2, y_2 lies to the left of γ_2. For this case let σ be the saddle connection of γ_2 that crosses R_4. We claim that

$$\operatorname{slope}(y_2y_3) \geq \operatorname{slope}(\sigma).$$

To see this, we focus attention on the rectangle R_4. By construction, γ_2 crosses R_4 along its top and right-hand edges. Also, y_2 is the bottom-right corner of R_4 and y_3 lies along the top edge of R_4 to the right of γ_2 (by the way Case 3 is defined). The claim follows.

We have the following path in $\mathcal{G}_{\widehat{X}}$:

$$R, y_1y_2, R_4, y_2y_3.$$

Applying the claim, the lemma follows in this case. □

4.2. Small-slope geodesics and approximating rectangles

The second statement of Lemma 4.1 is a consequence of the first part of the lemma and the following auxiliary lemma, which uses the tools of Section 3.

Lemma 4.2. *Let X be a singular Euclidean surface, let R be a maximal rectangle in $\widehat{X}$, let $F = 2\operatorname{flex}(X) - 1$, let $K = K(X)$ be the separatrix complexity, and let Z be a collection of pairwise noncrossing saddle connections in X. Assume that*

$$\min \operatorname{slope}(Z) < F^{-(2K+1)} \operatorname{slope}(R).$$

Then there is a $\sigma \in Z$ and a lift $\widetilde{\sigma} \subseteq \widehat{X}$ that crosses the vertical edges of R.

Proof. It suffices to consider the case where Z is a singleton $\{\sigma\}$. Let $\widetilde{\sigma}_-$ be the saddle connection in R of smallest slope, that is, the one joining its vertical edges, and let σ_- be its image in X. Let R_- be the rectangle spanned by $\widetilde{\sigma}_-$. The vertical edges of R_- are contained in those of R.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

By Lemma 3.5 there is a saddle connection ρ in $\mathbf{r}(\sigma)$ that has slope bounded above by $\operatorname{slope}(\sigma)$. In particular, $\operatorname{slope}(\rho) \leq F^{-(2K+1)} \operatorname{slope}(R)$. By Lemma 3.3, $\operatorname{slope}(\rho) \leq F^{-2K} \operatorname{slope}(\sigma_-)$. Then by Lemma 3.4, ρ crosses σ_-.

Let $\widetilde{\rho}$ be a lift of ρ to $\widehat{X}$ that crosses $\widetilde{\sigma}_-$. Since (as in the previous paragraph) $\operatorname{slope}(\rho) \leq \operatorname{slope}(\sigma_-)$, since $\widetilde{\rho}$ spans a rectangle, it must be that $\widetilde{\rho}$ crosses the vertical edges of R_-.

Choose a lift $\widetilde{\sigma}$ of σ with $\widetilde{\rho} \in \mathbf{r}(\widetilde{\sigma})$. By the definition of $\mathbf{r}(\widetilde{\sigma})$, there is a rectangle $R_{\widetilde{\rho}}$ in $\widehat{X}$ that contains $\widetilde{\rho}$ and has a segment of $\widetilde{\sigma}$ as its diagonal.

Since $R_{\widetilde{\rho}}$ contains no singularities in its interior and $\widetilde{\rho}$ crosses the vertical edges of R_-, the diagonal of $R_{\widetilde{\rho}}$ must also cross the vertical edges of R_-. But this diagonal is contained in $\widetilde{\sigma}$ and so $\widetilde{\sigma}$ crosses the vertical edges of R_- and hence of R. This completes the proof. □

Proof of Lemma 4.1(2). By the assumptions in the statement and the definition of a k-approximating rectangle we have

$$\operatorname{slope}(Z) < F^{-(2K+6)} \min\{\operatorname{slope}(\gamma_1) \cup \operatorname{slope}(\gamma_2)\} \leq F^{-(2K+1)} \operatorname{slope}(R).$$

An application of Lemma 4.2 completes the proof. □

5. The forcing lemma

The goal of the section is to prove Proposition 5.1, which we call the forcing lemma. Roughly, it tells us that if we have a pseudo-Anosov mapping class f (hence a notion of slope) and a train track τ, then τ carries the curves of very small slope if and only if it carries the horizontal foliation for f.

For the statement, we recall some basic definitions regarding train tracks. Let τ be a train track in a surface S. We say that τ is filling if each of its complementary regions in S is a disk with at most one puncture; each such region can be regarded as a polygon. An extension of τ is a train track τ' with τ a subtrack. If τ' is obtained from τ by adjoining diagonals of complementary polygons, then we say that τ' is a diagonal extension. Next, we say that τ is transversely recurrent if for every branch b of τ there is a curve c in S so that c is transverse to b and forms no bigons with τ. As in the introduction, if c is a curve (or foliation) carried by τ then $\tau[c]$ is the sub-train track whose branches are traversed by c. We say that c is fully carried by τ if $\tau[c] = \tau$.

Proposition 5.1. *Let $f \in \operatorname{Mod}(S)$ be pseudo-Anosov with horizontal foliation $\mathcal{F}_h$, and let τ be a filling, transversely recurrent train track in S. If c is a curve in S or a foliation in S with*

$$\mathcal{H}_f(c) \geq \mathcal{H}_f(\tau) + Q_{\text{force}},$$

then the following hold.

1. *If $c \prec \tau$, then $\mathcal{F}_h \prec \tau[c]$.*
2. *If $\mathcal{F}_h$ is fully carried by τ, then c is fully carried by a diagonal extension of τ.*

We begin in Section 5.1 by defining left- and right-turning paths and explaining some of their properties. We then proceed to define the slope of a train track and relate it to the slopes of the left- and right-turning paths (Lemma 5.3). In Section 5.2 we introduce the notion of a shadow for the lift of a train track to the universal cover of the ambient surface. We then use the notion of a shadow to define obstructing branches and agreeable branches, and use these to give criterion for a foliation to be carried on a train track and a criterion for a carried foliation to traverse a branch of a train track, respectively. In Section 5.3 we prove Proposition 5.1 by first proving a version for train track branches, Lemma 5.6. The latter is an application of the approximating rectangles lemma, Lemma 4.1.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

5.1. Left- and right-turning paths and slopes of train tracks

The goal of this section is to define left- and right-turning paths for a lifted train track, and to compare their slopes to those of the corresponding train track. We begin with preliminaries from the theory of train tracks: transverse recurrence and quasi-geodesic train tracks.

Transverse recurrence and quasi-geodesic train tracks. If τ is a transversely recurrent train track in a hyperbolic surface S, then the preimage $\widetilde{\tau}$ of τ in $\mathbb{H}^2$ is quasi-geodesic, in the sense that every bi-infinite train path in $\widetilde{\tau}$ is a quasi-geodesic with uniform quasi-isometry constants that do not depend on the path [25]. This fact is the main reason why we frequently employ the assumption on transverse recurrence.

Because of this, we may apply the tools and techniques from hyperbolic geometry to $\widetilde{\tau}$. For instance, every bi-infinite train path γ in $\widetilde{\tau}$ has well-defined endpoints in $\partial\mathbb{H}^2$. For a singular Euclidean surface X there is a geodesic in the completed universal cover $\widehat{X}$ with the same endpoints as γ (this geodesic is unique except when γ corresponds to the lift of the core of a flat annulus). We will use the fact that two such paths γ_1 and γ_2 are linked at infinity if and only if their geodesic representatives in $\widehat{X}$ are.

As in the work of Masur–Minsky [20, Lemma 4.5], a transversely recurrent track $\widetilde{\tau} \subset \mathbb{H}^2$ does not admit any generalized bigons, by which we mean a region in $\mathbb{H}^2$ bounded by either *(i)* distinct, bi-infinite train paths with common endpoints in $\partial\mathbb{H}^2$ or *(ii)* distinct rays with common starting points in $\mathbb{H}^2$ and common ending points in $\partial\mathbb{H}^2$.

Left- and right-turning paths and cycles. Let τ be a train track in either a hyperbolic surface S or in the universal cover $\mathbb{H}^2$. Let b be a branch of τ. The *left-turning path* in τ through b is the bi-infinite train path γ_b^L in τ that contains b and has the following property: if we travel along γ_b^L from b to another branch, then at each switch we traverse the left-most branch in the direction of travel.

When τ is a train track in a surface S and b is an oriented branch in τ, we may also define the *left-turning cycle* β_b^L via the following construction. We start from b and proceed in the direction of the orientation, always turning left, until we arrive at a branch b' that we have already traversed in the direction we are about to traverse it again. Then β_b^L is the closed train path obtained by discarding the initial path from b to b'. Right-turning paths and cycles γ_b^R and β_b^R are defined similarly.

We have the following basic facts about a transversely recurrent train track τ in a surface S:

- for a branch b of τ the paths γ_b^L and γ_b^R are simple,
- for each oriented branch b of τ the curves β_b^L and β_b^R are simple, and
- for any oriented edge b, the paths and curves γ_b^L and β_b^L are disjoint, and similarly for γ_b^R and β_b^R.

All three statements are verified by checking that the lifts to $\mathbb{H}^2$ can be realized disjointly, equivalently that there are no two (components of) lifts whose endpoints separate each other in $\partial\mathbb{H}^2$. If this were the case, then at the first and last intersections of the first path with the second, the first path would be turning in the same direction with respect to the second, which is impossible.

One other fact that we will use is that each curve of the form β_b^L or β_b^R is a vertex cycle for τ. This is straightforward from the definition; see, for example, [14, Lemma 2.12].

Slopes of train tracks versus slopes of left- and right-turning paths. We define the slope of a train track τ to be the union of the slopes of its vertex cycles, and we define $\mathcal{H}_f(\tau)$ in the usual way from the slope. Recall that $Q_{\text{curve}} = 2K^2$, that $E = 14$ and that $Q_{\text{track}} = EQ_{\text{curve}}$. We prove now two lemmas in turn.

Lemma 5.2. *Let τ be a transversely recurrent train track in a surface S. Then*

$$\operatorname{diam} \mathcal{H}_f(\tau) \leq Q_{\text{track}}.$$

Proof. As in the introduction, $E(S)$ is an upper bound for the diameter in $\mathcal{C}(S)$ of the set of vertex cycles of τ. By Proposition 3.1, each edge (v, w) in $\mathcal{C}(S)$ has $\operatorname{diam} \mathcal{H}_f(v) \cup \mathcal{H}_f(w) \leq Q_{\text{curve}}$. Finally, for any edge in the curve graph, the corresponding saddle connections are pairwise noncrossing (this is even true in the exceptional cases of $S_{1,1}$ and $S_{0,4}$ where edges do not correspond to disjoint curves but to curves that intersect once and twice, respectively). The lemma follows. □

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

Lemma 5.3. *Let $f \in \mathrm{Mod}(S)$ be pseudo-Anosov with singular Euclidean surface X and let $K = K(X)$. If $\tau \subseteq X$ is a transversely recurrent train track then for each branch b of τ we have*

$$\mathcal{H}_f(\tau) - (Q_{\text{track}} + Q_{\text{curve}}) \leq \mathcal{H}_f(\gamma_b^L) \leq \mathcal{H}_f(\tau) + (Q_{\text{track}} + Q_{\text{curve}})$$

and similarly for γ_b^R.

Proof. As above, we have that every γ_b^L is disjoint from the corresponding vertex cycle β_b^L. Again by Proposition 3.1 we have $\operatorname{diam} \mathcal{H}_f(\gamma_b^L) \cup \mathcal{H}_f(\beta_b^L) \leq Q_{\text{curve}}$. By Lemma 5.2 and the fact that β_b^L is a vertex cycle of τ, the lemma follows. □

5.2. Obstructing and agreeable branches

In this section, we introduce the notion of a shadow for a train track. We then define the notion of an obstructing branch and use it to give a criterion for a foliation to be carried by a train track. Finally, we define the notion of an agreeable branch and use it to give a criterion for a carried foliation to traverse a certain branch of a train track.

Shadows, obstructing branches, and agreeable branches. Let τ be a transversely recurrent train track in a hyperbolic surface S and $\widetilde{\tau}$ the preimage in $\mathbb{H}^2$. Let b be a branch of $\widetilde{\tau}$. We define the shadow Λ_b as the union of the two disjoint intervals of $\partial\mathbb{H}^2$ that run counterclockwise from an endpoint of γ_b^R to an endpoint of γ_b^L.

Next, let γ be a bi-infinite path in $\mathbb{H}^2$ with well-defined endpoints at infinity (for instance a component of the preimage of a curve in S or a leaf of a foliation of S). We say that b obstructs γ if the endpoints of γ in $\partial\mathbb{H}^2$ separate the two components of the shadow Λ_b (in the language of Masur–Minsky [20, Proof of Lemma 4.5], b consistently separates γ). We then say that b obstructs a foliation $\mathcal{F}$ of S if it is obstructing for some component of the preimage of some singularity-free leaf of $\mathcal{F}$.

Finally, we say that the branch b is agreeable to the path γ if the endpoints of γ in $\partial\mathbb{H}^2$ lie in different components of Λ_b. As in the obstructing case, we say that b is agreeable to a foliation $\mathcal{F}$ if it is agreeable to some component of the preimage of some singularity-free leaf of $\mathcal{F}$.

Obstructing branches and a criterion for carrying. The next lemma uses obstructing branches to give a criterion for a curve or a foliation to be carried by a train track.

Lemma 5.4. *Let τ be a filling transversely recurrent train track in a hyperbolic surface S, and let $\mathcal{F}$ be a measured foliation of S. Then $\mathcal{F}$ is carried on a diagonal extension of τ if and only if τ has no obstructing branches for $\mathcal{F}$.*

Proof. A proof of the reverse implication can be found in the work of Masur–Minsky [20, Proof of Lemma 4.5]. The argument there is framed in the special case where $\mathcal{F}$ is a simple closed curve (as opposed to a general foliation) but the same argument applies. (The hypotheses of the Masur–Minsky lemma include the assumption that τ is bi-recurrent; however, this part of the proof only uses the transverse recurrence).

It remains to prove the forward implication. Suppose that $\mathcal{F}$ is carried by τ. Let ℓ be a singularity-free leaf of the preimage of $\mathcal{F}$ in $\mathbb{H}^2$. Also let $\widetilde{\tau}$ be the preimage of τ in $\mathbb{H}^2$ and let b be a branch of $\widetilde{\tau}$. We would like to show that b is not an obstructing branch for ℓ.

Let γ be the bi-infinite train path in $\widetilde{\tau}$ whose endpoints in $\partial\mathbb{H}^2$ are the same as those of ℓ. Consider now the intersection $\gamma \cap \gamma_b^L$. If this intersection is empty, then the endpoints of ℓ and γ_b^L do not link in $\partial\mathbb{H}^2$ and it follows that b is not obstructing for ℓ.

Suppose then that b' is a branch of $\gamma \cap \gamma_b^L$. There are two subpaths of γ that start at b'. Choose a subpath γ' that moves away from b (if $b = b'$, then both choices move away from b). Since b' lies on γ_b^L, the endpoint of γ' at infinity lies in the shadow Λ_b. Thus b is not obstructing for ℓ, as desired. □

Agreeable branches and a criterion for edge traversal. The next lemma uses agreeable branches to give a criterion for a carried curve or foliation to traverse a particular edge of a train track.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

Lemma 5.5. *Let $\tau \subset S$ be a transversely recurrent train track in a hyperbolic surface, let $\widetilde{\tau}$ be the preimage in $\mathbb{H}^2$, and let γ be a bi-infinite train path in $\widetilde{\tau}$. Suppose that the endpoints of γ in $\partial\mathbb{H}^2$ lie in different components of Λ_b for some branch b of $\widetilde{\tau}$. Then γ traverses b.*

Proof. Let α be the intersection of γ_b^L and γ_b^R. It must be that α is a connected train path in $\widetilde{\tau}$ (otherwise $\widetilde{\tau}$, hence τ, would bound an embedded bigon). So $\gamma_b^L \cup \gamma_b^R$ divides $\mathbb{H}^2$ into four regions, two that meet Λ_b and two others. If γ does not traverse α, then it must enter the interior of one of the latter two regions. The intersection of γ with this region is a train path that is either

- a compact train path,
- an infinite train path asymptotic to $\partial\Lambda_b$, or
- a bi-infinite train path (namely, all of γ) asymptotic to $\partial\Lambda_b$ at both ends.

In the first case, we see that $\widetilde{\tau}$, hence τ, bounds a bigon, and in the second and third cases we see that $\widetilde{\tau}$ has a generalized bigon. These contradictions complete the proof. □

5.3. *Proof of the forcing lemma*

In this section, we prove Proposition 5.1, the forcing lemma. This proposition concludes that certain curves or leaves of a foliation are carried by a train track and also that every branch of the train track is traversed by this carrying. Before proceeding to the proof of the forcing lemma, we give a local version, that is, a version for a single branch of a (lifted) train track.

Lemma 5.6. *Let $f \in \mathrm{Mod}(S)$ be pseudo-Anosov with singular Euclidean surface X and horizontal foliation $\mathcal{F}_h$. Let $\tau \subset S$ be a filling transversely recurrent train track and let η be a curve or a leaf of a foliation in S with*

$$\mathcal{H}_f(\eta) \geq \mathcal{H}_f(\tau) + Q_{\mathrm{force}}.$$

Let $\widetilde{\tau}$, $\widetilde{\mathcal{F}_h}$, and $\widetilde{\eta}$ be the preimages in $\mathbb{H}^2$ of τ, $\mathcal{F}_h$, and η. Let b be a branch of $\widetilde{\tau}$, and let ℓ be a singularity-free leaf of $\widetilde{\mathcal{F}_h}$.

1. *If b is agreeable to ℓ then b is agreeable to some component of $\widetilde{\eta}$.*
2. *If b is obstructing for ℓ then b is obstructing for some component of $\widetilde{\eta}$.*

Proof. We would like to apply Lemma 4.1 to the triple $(\gamma_b^L, \gamma_b^R, \ell)$, where the paths γ_b^L and γ_b^R have been replaced with their geodesic representatives in $\widehat{X}$. To this end, we verify the hypotheses of that lemma. We have assumed in the hypotheses of this lemma that $\mathrm{flex}(X) < \infty$ and that ℓ contains no singularities. We also have that γ_b^L and γ_b^R intersect, since they are the geodesic representatives of left- and right-turning paths through b (and hence must be linked at infinity).

We claim that ℓ crosses both γ_b^L and γ_b^R. Indeed, the only other possibility is that an endpoint of ℓ meets Λ_b in its boundary. This is impossible because the points of $\partial\Lambda_b$ are endpoints of lifts of the vertex cycles of τ corresponding to the left- and right-turning cycles through the image of b.

Since the hypotheses of Lemma 4.1 are satisfied, it follows from the first statement of the lemma that there is an F^5-approximating rectangle R for the triple $(\gamma_b^L, \gamma_b^R, \ell)$ where F denotes the quantity $2\,\mathrm{flex}(X) - 1$.

We now would like to apply the second statement of Lemma 4.1 to η, which requires us to prove

$$\min \mathrm{slope}(\eta) < F^{-(2K+6)} \min\{\mathrm{slope}(\gamma_b^L) \cup \mathrm{slope}(\gamma_b^R)\}$$

By our hypothesis on η and Lemma 5.3, we have

$$\begin{aligned}\mathcal{H}_f(\eta) &\geq \mathcal{H}_f(\tau) + Q_{\mathrm{force}} \\ &= \mathcal{H}_f(\tau) + Q_{\mathrm{track}} + (2K+6)\cdot 2K \\ &\geq \max\{\mathcal{H}_f(\gamma_b^L) \cup \mathcal{H}_f(\gamma_b^R)\} + (2K+6)\cdot 2K\end{aligned}$$

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

Converting horizontalities to slope, we have

$$\operatorname{slope}(\eta) \leq (\lambda^{2K})^{-(2K+6)} \min\{\operatorname{slope}(\gamma_b^L) \cup \operatorname{slope}(\gamma_b^R)\}$$

Hence to verify the hypothesis for η it suffices to show that $F < \lambda^{2K}$. But this follows from Lemmas 3.6 and 3.7:

$$F = 2\operatorname{flex}(X) - 1 \leq 2\lambda^K - 1 < 2\lambda^K \leq \lambda^{2K}.$$

Having verified the hypothesis on η, Lemma 4.1(2) gives a geodesic $\widetilde{\eta}$ that is the geodesic representative of a component of the preimage of η and that crosses the vertical edges of R. Since R is an approximating rectangle, the geodesics γ_b^L and γ_b^R cross the horizontal edges of R. It follows that $\widetilde{\eta}$ crosses both γ_b^L and γ_b^R transversely within R.

In a CAT(0) space, the intersection of two geodesics is connected, and so $\widetilde{\eta}$ intersects each of γ_b^L and γ_b^R in a single point. Also, two geodesics that cross in a CAT(0) space cannot have the same endpoints at infinity. Therefore, the endpoints of $\widetilde{\eta}$ must be contained in the same components of $\partial\mathbb{H}^2 \setminus (\partial\gamma_b^L \bigcup \partial\gamma_b^R)$ as the endpoints of ℓ. This implies both statements of the lemma. □

Proof of Proposition 5.1. As in the statement, $\mathcal{F}_h$ is the horizontal foliation for a pseudo-Anosov $f \in \operatorname{Mod}(S)3$. Also, τ is a filling transversely recurrent train track in S, and c is a singularity-free leaf of a foliation in S with

$$\mathcal{H}_f(c) \geq \mathcal{H}_f(\tau) + Q_{\text{force}}.$$

In what follows, X is a singular Euclidean structure associated to f, and $\widetilde{\tau}$ and $\widetilde{\mathcal{F}_h}$ are the preimages of τ and $\mathcal{F}_h$ in the universal cover $\widetilde{X}$. We prove the two statements of the proposition in turn.

Statement (1). Assume that $c \prec \tau$. We first claim that $\mathcal{F}_h$ is carried by a diagonal extension τ' of τ. Assume not. By Lemma 5.4 there is a leaf ℓ of $\widetilde{\mathcal{F}_h}$ and a branch b of $\widetilde{\tau}$ that is obstructing for ℓ. Then by Lemma 5.6, c is obstructed by b. By Lemma 5.4, c is not carried by τ, the desired contradiction.

We now show that $\mathcal{F}_h$ is carried by $\tau'[c]$. Since $\tau'[c] = \tau[c]$, this will complete the proof of the first statement. Let b be a branch of $\widetilde{\tau}'$ traversed by some singularity-free leaf ℓ of $\widetilde{\mathcal{F}_h}$. The branch b is agreeable to ℓ. By Lemma 5.6, the branch b is agreeable to some component $\widetilde{c}$ of the preimage of c. By Lemma 5.5, the train path of $\widetilde{\tau}'$ corresponding to $\widetilde{c}$ traverses the branch b and the first statement follows.

Statement (2). As per the second statement, we assume that $\mathcal{F}_h$ is fully carried by τ. Let b be a branch of $\widetilde{\tau}$ traversed by (a leaf of) $\widetilde{\mathcal{F}_h}$. As in the proof of Statement (1), Lemma 5.6 implies there is a component of the preimage of c that traverses b. Since the components of $\widetilde{c}$ do not cross, it follows that b is not obstructing for any component of the preimage of c. Since b was arbitrary, Lemma 5.4 implies that c is carried by a diagonal extension τ' of τ. Applying Lemma 5.6 to τ', we conclude that c is fully carried by $\tau'[\mathcal{F}_h]$. Since the latter is equal to τ, the proof is complete. □

6. The fitting lemma

The goal of this section is to prove the following, which we call the fitting lemma.

Proposition 6.1. *Let $f \in \operatorname{Mod}(S)$ be pseudo-Anosov with horizontal foliation $\mathcal{F}_h$, and let $M \in \mathbb{R}$. Let T be the set of transversely recurrent train tracks in S that carry $\mathcal{F}_h$ and have horizontality at most M. There is an f-invariant train track ν so that ν is carried by each element of T and*

$$\mathcal{H}_f(\nu) \leq M + Q_{\text{fit}}.$$

The ν in the conclusion of Proposition 6.1 is fitted in the sense that it lies between $\mathcal{F}_h$ and each $\tau \in T$ (in the sense that $\mathcal{F}_h \prec \nu \prec \tau$) and it has horizontality close to $\mathcal{H}_f(T)$.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

The starting point for the proof of Proposition 6.1 is the work of Agol [1]. Given a pseudo-Anosov f, Agol gives a bi-infinite sequence of train tracks and multisplittings

$$\cdots \rightharpoonup \nu_i \rightharpoonup \nu_{i+1} \rightharpoonup \cdots \rightharpoonup \nu_{i+k} \rightharpoonup \cdots$$

and a $k > 0$ so that

$$f(\nu_i) = \nu_{i+k}$$

for all i; where by a multisplitting we mean a collection of simultaneous train track splits along one or more large branches. In particular, each ν_i is f-invariant and the $\mathcal{H}_f(\nu_i)$ tend to $\pm\infty$ as $i \to \pm\infty$. Also, each ν_i fully carries the unstable foliation for f, and the set of large branches split at each step is precisely the set of branches of largest weight with respect to the unstable foliation. The train track ν in the conclusion of Proposition 6.1 will be chosen from among the ν_i in Agol's sequence.

Agol proves that if τ is a train track carrying the horizontal foliation for f, then τ splits to some ν_i in the sequence [1, Theorem 3.5]. The basic idea of the proof of Proposition 6.1 is to use this fact to find a ν_i that all the τ_i split to, and then to show that we can move backwards along the sequence of ν_i until the slope condition is also satisfied.

We divide the proof into two steps. In Section 6.1 we prove Lemma 6.4, the backfolding lemma, which gives conditions under which the fold of a train track carried by a train track τ is again carried by τ; this is the main technical lemma of the section. Then in Section 6.2 we focus specifically on the Agol train tracks described above, and use them to prove Proposition 6.1.

6.1. The backfolding lemma

Before proceeding to the statement and the proof of the backfolding lemma, we introduce the following criterion for when a single fold does not disturb the carrying relation. In the statement, we say that a shadow Λ_b is contained in a shadow Λ_a by a component-respecting inclusion if each component of Λ_a contains one component of Λ_b.

Lemma 6.2. *Let S be a hyperbolic surface and let σ and τ be transversely recurrent train tracks in S with $\sigma \prec \tau$. Let $\sigma \mapsto \sigma'$ be a single fold creating a new large branch b'. Let $\widetilde{\sigma}'$ and $\widetilde{\tau}$ be the preimages of σ' and τ in $\mathbb{H}^2$, and let $\widetilde{b}' \subseteq \widetilde{\sigma}'$ be a lift of b' to $\mathbb{H}^2$. Suppose there is a branch $\tilde{a}$ of $\widetilde{\tau}$ with*

$$\Lambda_{\widetilde{b}'} \subseteq \Lambda_{\widetilde{a}}$$

with a component-respecting inclusion. Then σ' is carried by τ.

Proof. Let e, f, and g be the branches that are folded together in σ to produce the branch b' of σ', where f is the small branch. There are two cases: we can either choose the carrying map $\sigma \to \tau$ so that a small neighborhood of f in σ maps homeomorphically or we cannot (here we are referring only to the carrying map, not the map to a tie neighborhood). In the latter case $\sigma \prec \tau$ factors through $\sigma \mapsto \sigma'$, contrary to assumption. So we may assume that a small neighborhood of f does map homeomorphically; say that the branch in the image is h.

One way to visualize the statements in the preceding paragraph is to draw a tie neighborhood N of τ and to consider $\sigma \subset N$ so that collapsing along the fibers give a carrying map $\sigma \prec \tau$. If some train track tie of N passes through all three branches e, f, and g, the carrying map $\sigma \prec \tau$ factors through $\sigma \mapsto \sigma'$. The only other possibility is that e, f, and g are configured in N as in Figure 9.

Let $\widetilde{e}$, $\widetilde{f}$, and $\widetilde{g}$ be preimages of e, f, and g in $\widetilde{\sigma}$ that are folded to create $\widetilde{b}'$. Because of the folding relation, we have

$$\Lambda_{\widetilde{e}} \cup \Lambda_{\widetilde{f}} \cup \Lambda_{\widetilde{g}} \subseteq \Lambda_{\widetilde{b}'}$$

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

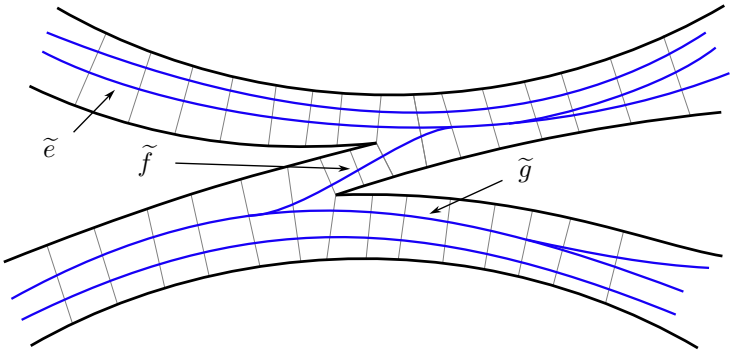


Figure 9. *The configuration of $\sigma \subset N$ if $\sigma \prec \tau$ does not factor through $\sigma \mapsto \sigma'$.*

by component-respecting inclusions. By assumption we have an edge $\tilde{a}$ of $\widetilde{\tau}$ with $\Lambda_{\widetilde{b}'} \subseteq \Lambda_{\widetilde{a}}$ and hence $\Lambda_{\widetilde{a}}$ contains $\Lambda_{\widetilde{e}}$, $\Lambda_{\widetilde{f}}$, and $\Lambda_{\widetilde{g}}$. Let $\widetilde{h}$ be the branch of $\widetilde{\tau}$ that is the homeomorphic image of $\widetilde{f}$.

By the above (component-respecting) inclusions, the path

$$\gamma = \gamma^L_{\widetilde{f}}$$

has endpoints in both components of $\Lambda_{\widetilde{a}}$. By Lemma 5.5, the image of γ in $\widetilde{\tau}$ passes through $\widetilde{a}$. In particular, $\widetilde{h}$ and the branch $\widetilde{a}$ are connected by a train path in $\widetilde{\tau}$ (a subpath of γ).

Finally, observe from Figure 9 that $\Lambda_{\widetilde{h}}$ contains neither $\Lambda_{\widetilde{e}}$ nor $\Lambda_{\widetilde{g}}$. Since moving along a train path does not increase the shadow in the forward direction, we must have that one of the inclusions $\Lambda_{\widetilde{e}} \subseteq \Lambda_{\widetilde{a}}$ or $\Lambda_{\widetilde{g}} \subseteq \Lambda_{\widetilde{a}}$ is false. This is the desired contradiction. □

We require one more lemma about slopes before proceeding to the backfolding lemma. Recall that $E = 14$ and $Q_{\text{curve}}(S) = 2K^2$.

Lemma 6.3. *Let $f \in \text{Mod}(S)$ be pseudo-Anosov, let $\sigma \mapsto \sigma'$ be a single fold of transversely recurrent train tracks, and let $b' \subseteq \sigma'$ be the large branch created in the fold. Then*

$$\min \mathcal{H}_f(\sigma) - (E+3)Q_{\text{curve}} \leq \mathcal{H}_f(\gamma^{L/R}_{b'}) \leq \max \mathcal{H}_f(\sigma) + (E+3)Q_{\text{curve}}.$$

Proof. By symmetry, it suffices to prove the first inequality. Let b be the branch of σ used to do the fold. Depending upon the orientation of the fold, it is either the case that $\gamma^L_{b'}$ and γ^L_b are homotopic or $\gamma^R_{b'}$ and γ^R_b are (up to modifying the fold itself by an ambient isotopy, we can in fact make the two paths equal). Say it is the former that are homotopic. As in the proof of Lemma 5.3, we have

$$\mathcal{H}_f(\gamma^L_{b'}) = \mathcal{H}_f(\gamma^L_b) \geq \min \mathcal{H}_f(\sigma) - Q_{\text{curve}}.$$

Using the same reasoning and applying Lemma 5.3, we have that

$$\mathcal{H}_f(\gamma^R_{b'}) \geq \mathcal{H}_f(\gamma^L_{b'}) - (E+2)Q_{\text{curve}} \geq \min \mathcal{H}_f(\sigma) - (E+2)Q_{\text{curve}}.$$

Combining these gives the lemma. □

We now arrive at the backfolding lemma.

Lemma 6.4. *Let $f \in \text{Mod}(S)$ be pseudo-Anosov with singular Euclidean surface X and horizontal and vertical foliations $\mathcal{F}_h$ and $\mathcal{F}_v$. Let τ be a transversely recurrent train track in S that fully carries $\mathcal{F}_h$, and let σ be a train track in S with $\sigma \prec \tau$ and*

$$\mathcal{H}_f(\sigma) \geq \mathcal{H}_f(\tau) + Q_{\text{fit}}$$

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

Further suppose that $\sigma \mapsto \sigma'$ is a single fold producing a large branch $b' \subseteq \sigma'$ and that b' obstructs $\mathcal{F}_v$. Then σ' is carried by τ.

Proof. Our goal is to apply Lemma 6.2. As such, our main work is to verify the hypotheses of that lemma, with τ, σ, σ', and b' in this lemma corresponding to the objects with the same names in the former. For starters, since f is pseudo-Anosov, we have that S is hyperbolic, as required by Lemma 6.2.

Let $\widetilde{b}'$ be a lift of b' to $\mathbb{H}^2$. The only hypothesis of Lemma 6.2 that it remains to verify is that there is a branch $\widetilde{a}$ in $\widetilde{\tau}$, the preimage of τ in $\mathbb{H}^2$, with

$$\Lambda_{\widetilde{b}'} \subseteq \Lambda_{\widetilde{a}}.$$

We first address the case where τ is complete, meaning that τ has no proper extensions, and then we use this special case to derive the general case.

Complete case. We first claim that $\gamma^L_{b'}$ and $\gamma^R_{b'}$ are carried by τ. We will prove this using Proposition 5.1, and so we would like to bound the horizontalities of both paths. First, the assumption $\mathcal{H}_f(\sigma) \geq \mathcal{H}_f(\tau) + Q_{\text{fit}}$ can be rephrased as $\min \mathcal{H}_f(\sigma) \geq \mathcal{H}_f(\tau) + Q_{\text{fit}}$. Recall that $Q_{\text{fit}} = Q_{\text{force}} + (E+3)Q_{\text{curve}}$. Using Lemma 6.3 and the rephrasing of the assumption we have

$$\begin{aligned}\mathcal{H}_f(\gamma^{L/R}_{b'}) &\geq \min \mathcal{H}_f(\sigma) - (E+3)Q_{\text{curve}} \\ &= \min \mathcal{H}_f(\sigma) - (Q_{\text{fit}} - Q_{\text{force}}) \\ &\geq \mathcal{H}_f(\tau) + Q_{\text{force}}.\end{aligned}$$

By Proposition 5.1, the assumption that τ fully carries $\mathcal{F}_h$ and the assumption that τ is complete, the claim now follows.

By the previous claim the paths $\gamma^L_{\widetilde{b}'}$ and $\gamma^R_{\widetilde{b}'}$ in $\mathbb{H}^2$ are carried by $\widetilde{\tau}$. Also, these two paths are linked at infinity in $\mathbb{H}^2$. Thus, there is a branch $\widetilde{a}$ in $\widetilde{\tau}$ contained in both paths. We will show that this $\widetilde{a}$ satisfies the desired containment $\Lambda_{\widetilde{b}'} \subseteq \Lambda_{\widetilde{a}}$.

Since $\gamma^L_{\widetilde{b}'}$ and $\gamma^R_{\widetilde{b}'}$ both pass through $\widetilde{a}$, we have

$$\partial \Lambda_{\widetilde{b}'} \subseteq \Lambda_{\widetilde{a}}.$$

From this containment, there are two possibilities for $\Lambda_{\widetilde{b}'}$: either $\Lambda_{\widetilde{b}'} \subseteq \Lambda_{\widetilde{a}}$ (which is what we want to show) or $\partial \Lambda_{\widetilde{a}} \subseteq \Lambda_{\widetilde{b}'}$. As such, it remains to rule out the latter.

By assumption there is a leaf ℓ of $\mathcal{F}_v$ that separates the components of $\Lambda_{\widetilde{b}'}$. Any such ℓ must cross both $\gamma^L_{\widetilde{b}'}$ and $\gamma^R_{\widetilde{b}'}$. By (the unstated vertical version of) Lemma 4.1 there is a maximal rectangle R so that ℓ meets R in a segment, so that the geodesic representative of $\gamma^L_{\widetilde{b}'}$ in $\widehat{X}$ crosses the vertical edges of R, and so that for every geodesic γ in X with

$$\max \text{slope}(\gamma) > F^{2K+6} \text{slope}(\gamma^i_{\widetilde{b}'}) \text{ for } i \in \{L, R\}$$

there is a component $\widetilde{\gamma}$ of the preimage of γ that crosses the horizontal edges of R. Here $F = 2\,\text{flex}(X) - 1$, as usual.

We next claim that $\gamma^L_{\widetilde{a}}$ satisfies the criterion

$$\max \text{slope}(\gamma^L_{\widetilde{a}}) > F^{2K+6} \text{slope}(\gamma^i_{\widetilde{b}'}) \text{ for } i \in \{L, R\}.$$

First, by Proposition 3.1, the assumption on $\mathcal{H}_f(\sigma)$, Lemma 6.3, the definition of E, and the fact that $Q_{\text{fit}} \geq 0$, we have

$$\begin{aligned}\min \mathcal{H}_f(\gamma^L_{\widetilde{a}}) &\leq \max \mathcal{H}_f(\tau) + Q_{\text{curve}} \\ &\leq \min \mathcal{H}_f(\sigma) + Q_{\text{curve}} - Q_{\text{fit}}\end{aligned}$$

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

$$\begin{aligned}&\leq \mathcal{H}_f(\gamma_{b'}^{L/R}) + (E+4)Q_{\text{curve}} - Q_{\text{fit}}\\&= \mathcal{H}_f(\gamma_{b'}^{L/R}) + 18\,Q_{\text{curve}} - Q_{\text{fit}}\\&\leq \mathcal{H}_f(\gamma_{b'}^{L/R}) + 18\,Q_{\text{curve}}.\end{aligned}$$

We convert to slopes and use Penner's inequality $\lambda^{2K} > F$ (Lemma 3.7):

$$\begin{aligned}\max \text{slope}(\gamma_{\widetilde{a}}^{L}) &\geq (\lambda^{2K})^{18K}\,\text{slope}(\gamma_{b'}^{L/R})\\&\geq F^{18K}\,\text{slope}(\gamma_{b'}^{L/R}).\end{aligned}$$

Since $18K > 2K+6$, the claim follows. (This argument does not use $\mathcal{H}_f(\sigma) \geq \mathcal{H}_f(\tau) + Q_{\text{fit}}$, but the weaker assumption that $\mathcal{H}_f(\sigma) \geq \mathcal{H}_f(\tau) + (E+2)Q_{\text{curve}} - 12K$.)

Let us replace $\gamma_{\widetilde{a}}^{L}$, $\gamma_{\widetilde{b}}^{L}$, and $\gamma_{\widetilde{b}}$ with their geodesic representatives in $\widehat{X}$. By the previous claim, the description of R from Lemma 4.1, and the CAT(0) geometry of $\widehat{X}$ (specifically that geodesics in $\widehat{X}$ intersect in a connected set), there is a translate γ' of $\gamma = \gamma_{\widetilde{a}}^{L}$ that crosses both $\gamma_{\widetilde{b}}^{L}$ and $\gamma_{\widetilde{b}}^{R}$ each in exactly one point. It follows that γ' also separates the components of $\Lambda_{\widetilde{b}'}$. However, under our assumption γ has its boundary in $\Lambda_{\widetilde{b}'}$ and so γ and γ' must cross. This contradicts the fact that left-turning paths are simple and we are done.

General case. Suppose that $\sigma' \rightharpoonup \sigma$ is a left split, so that γ_b^L is automatically carried by τ. Now by Proposition 5.1 and the assumption that τ fully carries $\mathcal{F}_h$, we have that $\gamma_{\widetilde{b}}^{R}$ is carried by $\widetilde{\tau}'$, where τ' is some diagonal extension of τ. Let $\widetilde{a}$ be a common branch of the train paths $\gamma_{\widetilde{b}}^{L/R}$ as carried by $\widetilde{\tau}'$. As $\gamma_{\widetilde{b}}^{L} \prec \widetilde{\tau}$, we must have that $\widetilde{a}$ is contained in $\widetilde{\tau}$. Note that since a diagonal of a diagonal extension is never the left/right-most turn, we have that $\Lambda_{\widetilde{a}}$ is independent of whether $\widetilde{a}$ is considered as a branch of $\widetilde{\tau}$ or $\widetilde{\tau}'$. By the special case of the lemma, we have $\Lambda_{\widetilde{b}} \subseteq \Lambda_{\widetilde{a}}$ and we are done. □

6.2. Proof of the fitting lemma

Before proving Proposition 6.1, we have a technical lemma about Agol's train tracks, Lemma 6.5. This lemma will use a fact from the work of Guéritaud, Minsky, and the third-named author, which we now explain.

Guéritaud and transversality. Using the work of Guéritaud, Minsky and the third-named author observed that each Agol train track ν_i can be realized as a smoothing of the dual graph to a triangulation of the surface by rectangle-spanning saddle connections. Moreover, the large half-branch in each triangle meets the tallest side. See Figure 10 for an illustration of the local picture in one such triangle. It follows that ν_i can be chosen to be transverse to the vertical foliation $\mathcal{F}_\nu$ for the corresponding pseudo-Anosov mapping class.

Lemma 6.5. *Let $f \in \text{Mod}(S)$ be pseudo-Anosov with vertical foliation $\mathcal{F}_\nu$, and let ν be an Agol train track for f. Then every branch of ν obstructs $\mathcal{F}_\nu$.*

Proof. Fix a singular Euclidean structure on S associated to f, and let $\widetilde{\nu}$ be the preimage of ν in the corresponding completed universal cover $\widehat{X}$. Let $\widetilde{\mathcal{F}_\nu}$ be the induced vertical foliation of $\widehat{X}$. As above, we may assume that $\widetilde{\nu}$ is transverse to $\widetilde{\mathcal{F}_\nu}$.

We claim that a train path in $\widetilde{\nu}$ meets a nonsingular leaf of $\widetilde{\mathcal{F}_\nu}$ at most once. Indeed, suppose that ℓ is a nonsingular leaf of $\widetilde{\mathcal{F}_\nu}$ that meets some train path of $\widetilde{\nu}$ in more than one point. There is an embedded (innermost) disk D in $\mathbb{H}^2$ bounded by a train path p in $\widetilde{\nu}$ and a segment of ℓ, and so that the interior of D does not meet $\widetilde{\nu}$. The disk D is foliated by (the restriction of) $\widetilde{\mathcal{F}_\nu}$. If we double D across the segment of ℓ, we obtain a foliated disk where the leaves are transverse to the boundary. Doubling this resulting disk along its boundary, we obtain a foliated sphere where each singularity corresponds to a singularity in the interior of D. Each such singularity has negative index, contradicting the Gauss–Bonnet formula. The claim follows.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

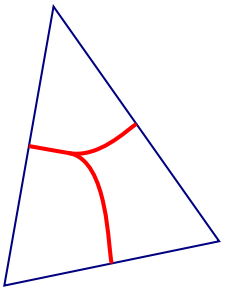


Figure 10. *A piece of an Agol train track inside a triangle of a triangulation by rectangle-spanning saddle connections.*

Now, let b be a branch of $\widetilde{\nu}$, and let ℓ be a nonsingular leaf of $\mathcal{F}_v$ that passes through b. By the claim, ℓ meets γ_b^L and γ_b^R exactly once each. Thus the endpoints of ℓ separate the components of Λ_b and so b obstructs ℓ, hence $\mathcal{F}_v$, as required. □

Proof of Proposition 6.1. As above, the $\mathcal{H}_f(\nu_i)$ tend to $\pm\infty$ as $i \to \pm\infty$. Thus there is an $i \in \mathbb{Z}$ so that $\mathcal{H}_f(\nu_i) \leq M + Q_{\text{fit}}$ and $\mathcal{H}_f(\nu_{i+k}) \geq M + Q_{\text{fit}}$ for all $k > 0$. To prove the proposition, it suffices to show that ν_i is carried by all elements of T.

So let $\tau \in T$. As above, this means that τ is a transversely recurrent train track in S that carries $\mathcal{F}_h$ and has $\mathcal{H}_f(\tau) \leq M$. Let τ' denote $\tau[\mathcal{F}_h]$, the sub-train track of τ fully carrying $\mathcal{F}_h$. We will show that $\nu_i \prec \tau'$.

By the work of Agol discussed at the start of the section, τ' splits to some ν_j in the Agol sequence. Assume that j is the smallest integer with $\nu_j \prec \tau'$. If $j \leq i$, we are done, because then $\nu_i \prec \nu_j \prec \tau'$. So we assume $j > i$. By our assumption in the first paragraph of the proof, $\mathcal{H}_f(\nu_j) \geq M + Q_{\text{fit}}$. Lemma 6.5 implies that every branch of ν_j obstructs the vertical foliation $\mathcal{F}_v$ for f.

Lemma 6.4 then implies that each individual fold of the multifold $\nu_j \mapsto \nu_{j-1}$ produces a train track carried by τ'. Since these individual folds are far apart from each other in the sense that the large branches they create are disjoint, this further implies that ν_{j-1} is carried by τ', contradicting the minimality of j. Thus, $j \leq i$ and so $\nu_i \prec \nu_j \prec \tau'$, as desired. □

7. Proof of Theorem 1.1

Before proving our main technical theorem, we require several technical lemmas: a criterion for carrying, a construction of a train track neighborhood of a curve, an upper bound on the radius of a Dehn–Thurston cell structure, an upper bound for the radius of the set of vertex cycles in a Dehn–Thurston train track, and a linear algebraic lemma about projective sinks.

Another carrying criterion. The following lemma gives a criterion for the extension operation to preserve the carrying relation between train tracks.

Lemma 7.1. *Suppose that $\nu \prec \tau$, where ν and τ are large train tracks in a surface. Let $\widetilde{\nu}$ be a diagonal extension of ν that fully carries a measured foliation y that is also carried by τ. Then a carrying map $\nu \to \tau$ extends to a carrying map $\widetilde{\nu} \to \tau$.*

Proof. A carrying map from ν to τ takes $\widetilde{\nu}$ to a diagonal extension $\widetilde{\tau}$ of τ (which we assume to be a minimal such extension). Since y is carried by τ and $\widetilde{\nu}$ and since the carrying position of y in τ is unique, we have $\widetilde{\tau} = \tau$, as desired. □

A train track neighborhood. We require the following standard construction of an open neighborhood of a point in $\mathrm{MF}(S)$.

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

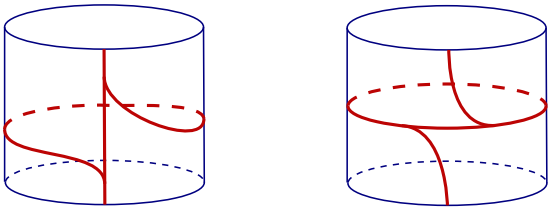


Figure 11. *Two model train tracks on an annulus carrying a common path connecting the boundaries.*

Lemma 7.2. *Suppose $x \in \mathrm{MF}(S)$ is fully carried by a filling train track σ, which is in turn carried by a train track τ. Let $\sigma_1, \dots, \sigma_k$ be the set of diagonal extensions of σ that are carried by τ. Then the union*

$$V(\sigma_1) \cup \cdots \cup V(\sigma_k)$$

contains an open neighborhood of x in $V(\tau)$.

Proof. The train track σ can be pinched along complementary components to produce a complete track $\widetilde{\sigma}$ with $\sigma \prec \widetilde{\sigma}$ and so that each curve fully carried by σ is fully carried by $\widetilde{\sigma}$; see [25, Proposition 1.4.9] where it is said that $\widetilde{\sigma}$ arises from σ by "trivial collapses along admissible arcs."

Say that $\sigma_1, \dots, \sigma_\ell$ are the diagonal extensions of σ (note $\ell \geq k$). The set $V(\widetilde{\sigma})$ is equal to the union $V(\sigma_1) \cup \cdots \cup V(\sigma_\ell)$ (see the "improvement" following Theorem 5.12 in the book by Casson and Bleiler [10]). The lemma follows. (Alternatively, the construction of the pinched complete track $\widetilde{\sigma}$ above can be done within a tie neighborhood of τ, yielding the same conclusion.) □

The radius for a Dehn–Thurston cell structure. The following lemma justifies our choice of $D(S) = 1$ in the introduction. Again, see the book by Penner–Harer for background on Dehn–Thurston train tracks [25, §2.6].

Lemma 7.3. *Let S be a surface, let Y be a Dehn–Thurston cell structure on $\mathrm{PMF}(S)$, and let c be a curve of the underlying pants decomposition. Then the radius of Y with respect to c is 1.*

Proof. There are two types of train tracks in the given collection of Dehn–Thurston train tracks: (1) those that contain c as a vertex cycle and (2) those that do not carry c. To prove the lemma we must show that each train track of the second type carries a curve in common with some train track of the first type.

So let τ be of the second type. This means that the model train track τ_A for τ in the annulus A around c does not carry c. Of the two model tracks on A that carry c, there is exactly one with the property that there is a homotopy class α of paths in A (relative to ∂A) carried both by this model and τ_A; see Figure 11. Let τ' be the train track obtained from τ by changing the model on A to this model.

Let d be a curve in S that is carried by τ and so that each component of the intersection with A is parallel to α (relative to ∂A); for instance, there is a vertex cycle for τ that passes through A once and satisfies this property. Then d is also carried by τ' since τ agrees with τ' outside A. This completes the proof. □

Sinks and eigenvalues. Finally, we give a lemma that allows us to conclude the second statement of Theorem 1.1 from the first. As in the introduction, a vertex ray for an integral cone in an integral cone complex Y is a vertex of the corresponding polyhedron in $\mathcal{P}(Y)$.

Lemma 7.4. *Suppose that Y is an integral cone complex and that the group G acts on Y with property Q. Let $f : Y \to Y$ be a projectively source-sink integral cone complex map with sink v, associated eigenvalue λ, and extended dynamical map given by (V_i, D). Suppose that v does not lie on a rational ray in V_i. Then the largest real eigenvalue of D is equal to λ and any corresponding positive eigenvector represents v.*

Proof of Lemma 7.4. As in the extended dynamical matrix algorithm, the extended dynamical matrix D induces a linear map of a subspace W of $\mathbb{R}^d$, and moreover v lies in W. Suppose that w is an eigenvector

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

of D with real eigenvalue $\mu > 1$. Then D acts on the subspace W' of W spanned by v and w. Also, by the description of D from the extended dynamical matrix algorithm, w lies in the subspace of $\mathbb{R}^d$ spanned by V_i.

Let $U \subseteq V_i$ be the PL-eigenregion used to determine D. Since v is not rational, it does not lie in a vertex ray of U, and so the intersection $W' \cap U$ contains a two-dimensional cone C (for the same reason that the intersection of a line with a polytope is either a vertex or contains an open interval).

The actions of f and D agree on C. Since f acts with projective source-sink dynamics on Y, we can choose a smaller cone C' around the span of v with the property that $f(C')$ is properly contained in C'. It follows that $\mu < \lambda$, as desired. The lemma follows. □

Proof of Theorem 1.1. We begin with the first statement. Let $\{\tau_i\}$ be a finite set of complete, transversely recurrent train tracks giving the integral cone complex structure Y on $\mathrm{MF}(S)$, and say that c is a vertex cycle of τ_0. Let C be a top-dimensional linear cone for the f-action on Y that contains $f^Q(c)$.

We will show that C shares an open subset with a Nielsen–Thurston eigenregion for f, namely, the region $V(\widetilde{\nu})$ below, and that both are mapped by f into the same $V(\tau_j)$. From this it follows that the two linear actions must be equal. In particular, there is a single f-linear piece containing both. This linear piece is still a Nielsen–Thurston eigenregion. The first statement of the theorem follows from this.

To fix notation, suppose that $C \subseteq V(\tau_i)$ and that f maps C linearly into $V(\tau_j)$ (while our proof applies in all cases, the proof simplifies in the case $i = j$, which is the generic case; we leave this simplification to the reader).

We claim that both $V(\tau_i)$ and $V(\tau_j)$ contain $\mathcal{F}_h$. Using the basic properties of horizontality, the definition of Q, and the definition of Q_{DT} with Proposition 3.1, we have:

$$\begin{aligned}\mathcal{H}_f(f^Q(c)) &= \mathcal{H}_f(c) + Q \\ &= \mathcal{H}_f(c) + Q_{\mathrm{force}} + Q_{\mathrm{fit}} + Q_{\mathrm{DT}} + 1 \\ &\geq \mathcal{H}_f(\tau_i) + Q_{\mathrm{force}} + Q_{\mathrm{fit}} + 1,\end{aligned}$$

and similarly for τ_j. As $Q_{\mathrm{force}} + Q_{\mathrm{fit}} + 1 > Q_{\mathrm{force}}$, the claim now follows from Proposition 5.1.

Let

$$M = \max \mathcal{H}_f(\tau_i) \cup \mathcal{H}_f(\tau_j).$$

By the previous claim, Proposition 6.1 now gives an f-invariant train track ν that fully carries $\mathcal{F}_h$, that is carried by both τ_i and τ_j, and that satisfies

$$\mathcal{H}_f(\nu) \leq M + Q_{\mathrm{fit}}.$$

We now claim that

$$\mathcal{H}_f(f^Q(c)) > \mathcal{H}_f(\nu) + Q_{\mathrm{force}}.$$

Indeed, let τ be the element of $\{\tau_i, \tau_j\}$ realizing M. Using the inequality obtained in the last claim, the definition of M, and the definition of ν we have

$$\begin{aligned}\mathcal{H}_f(f^Q(c)) &\geq \max \mathcal{H}_f(\tau) + Q_{\mathrm{force}} + Q_{\mathrm{fit}} + 1 \\ &= M + Q_{\mathrm{force}} + Q_{\mathrm{fit}} + 1 \\ &\geq \mathcal{H}_f(\nu) + Q_{\mathrm{force}} + 1,\end{aligned}$$

and the claim follows.

Combining the previous claim with Proposition 5.1, we obtain that $f^Q(c)$ is fully carried by some diagonal extension ν_1 of ν. As $f^Q(c)$ and ν are each carried by τ_i, it follows from Lemma 7.1 that ν_1 is also carried by τ_i. It then follows from Lemma 7.2 that there is a diagonal extension $\widetilde{\nu}$ of ν_1 that is

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press

carried by τ_i and so that the interiors of $V(\widetilde{\nu})$ and C have nonempty intersection ($\widetilde{\nu}$ is also an extension of ν, hence the notation). Let y be a point in this intersection. This means in particular that $\widetilde{\nu}$ fully carries y. As above, our goal is to show $V(\widetilde{\nu})$ is a Nielsen–Thurston eigenregion.

To this end, we claim $f(\widetilde{\nu}) \prec \tau_j$. First, we have $\nu \prec \tau_j$ so $f(\nu) \prec \tau_j$. Also, since $f(C)$ lies in $V(\tau_j)$ we have $f(y) \prec \tau_j$. Since $\widetilde{\nu}$ fully carries y, $f(\widetilde{\nu})$ fully carries $f(y)$, and so Lemma 7.1 implies that τ_j carries the unique diagonal extension of $f(\nu)$ that fully carries $f(y)$. But this unique extension must be $f(\widetilde{\nu})$, since $\widetilde{\nu}$ is the unique extension of ν fully carrying y and since f is (represented by) a homeomorphism.

We may now conclude that $V(\widetilde{\nu})$ is a Nielsen–Thurston eigenregion for f. Indeed, $V(\widetilde{\nu})$ contains $\mathcal{F}_h$ since $V(\nu)$ does, and the last claim implies that f is linear on $V(\widetilde{\nu})$. As discussed, this completes the proof of the first statement.

The second statement follows from the first statement and Lemma 7.4. Indeed, for the $\mathrm{Mod}(S)$-action on the Dehn–Thurston integral cone complex, the foliation $\mathcal{F}_h$ automatically satisfies the condition of not lying in a vertex ray. The reason is that in constructing the linear pieces for f as in the basic computational algorithm, the subdivisions of the cells of Y are made using rational planes. It follows that the vertices of the resulting subdivision of $\mathcal{P}(Y)$ are rational, which means they correspond to multicurves in S. On the other hand, the unstable foliation for a pseudo-Anosov mapping class does not correspond to a multicurve. This completes the proof. □

Competing interests. The authors have no competing interests to declare.

Funding statement. This material is based upon work supported by the National Science Foundation under Grant Nos. DMS-1057874, DMS-1811941, DMS-2102018, and DMS-2203431. We also acknowledge support through a Sloan Research Fellowship and a Fulbright Senior Scholar Award.

References

[1] I. Agol, 'Ideal triangulations of pseudo-Anosov mapping tori', in *Topology and Geometry in Dimension Three*, Contemp. Math. vol. 560 (Amer. Math. Soc., Providence, RI, 2011), 1–17.
[2] T. Aougab, 'Uniform hyperbolicity of the graphs of curves', *Geom. Topol.* **17**(5) (2013), 2855–2875.
[3] F. Baroni, 'Uniformly polynomial-time classification of surface homeomorphisms', 2024, https://link.springer.com/article/10.1007/s00208-025-03159-w.
[4] M. C. Bell and S. Schleimer, 'Slow north-south dynamics on $\mathcal{PML}$', *Groups Geom. Dyn.* **11**(3) (2017), 1103–1112.
[5] M. C. Bell and R. C. H. Webb, 'Polynomial-time algorithms for the curve graph', 2016, https://arxiv.org/abs/1609.09392.
[6] M. C. Bell, *Recognising Mapping Classes*, PhD thesis, University of Warwick, 2015.
[7] M. Bestvina and M. Handel, 'Train-tracks for surface homeomorphisms', *Topology* **34**(1) (1995), 109–140.
[8] B. H. Bowditch, 'Uniform hyperbolicity of the curve graphs', *Pacific J. Math.* **269**(2) (2014), 269–280.
[9] P. Brinkmann, 'An implementation of the Bestvina-Handel algorithm for surface homeomorphisms', *Experiment. Math.* **9**(2) (2000), 235–240.
[10] A. J. Casson and S. A. Bleiler, Automorphisms of surfaces. *Notes.*
[11] B. Farb and D. Margalit, *A Primer on Mapping Class Groups*, Princeton Mathematical Series vol. 49 (Princeton University Press, Princeton, NJ, 2012).
[12] A. Fathi, F. Laudenbach and V. Poénaru, *Travaux de Thurston sur les surfaces*, Astérisque vol. 66 (Société Mathématique de France, Paris, 1979). Séminaire Orsay, With an English summary.
[13] M. D. Finn and J.-L. Thiffeault, 'Topological entropy of braids on the torus', *SIAM J. Appl. Dyn. Syst.* **6**(1) (2007), 79–98 (electronic).
[14] V. Gadre and S. Schleimer, 'The curves not carried', *Groups Geom. Dyn.* **10**(4) (2016), 1249–1264.
[15] V. Gadre and C.-Y. Tsai, 'Minimal pseudo-Anosov translation lengths on the complex of curves', *Geom. Topol.* **15**(3) (2011), 1297–1312.
[16] T. Hall, Dynn (computer software), Version 0.10.3, 2015.
[17] T. Hall and S. Ö. Yurttaş, 'On the topological entropy of families of braids', *Topology Appl.* **156**(8) (2009), 1554–1564.
[18] H. Hamidi-Tehrani and Z.-H. Chen, 'Surface diffeomorphisms via train-tracks', *Topology Appl.* **73**(2) (1996), 141–167.
[19] W. Jaco and J. L. Tollefson, 'Algorithms for the complete decomposition of a closed 3-manifold', *Illinois J. Math.* **39**(3) (1995), 358–406.
[20] H. A. Masur and Y. N. Minsky, 'Geometry of the complex of curves. I. Hyperbolicity', *Invent. Math.* **138**(1) (1999), 103–149.
[21] C. Menzel and J. R. Parker, 'Pseudo-Anosov diffeomorphisms of the twice-punctured torus', in J. R. Cho and J. Mennicke, editors, *Recent Advances in Group Theory and Low-Dimensional Topology*, no. 27 (Heldermann Verlag, 2003), 141–154. Unpublished.

[22] Y. N. Minsky and S. J. Taylor, 'Fibered faces, veering triangulations, and the arc complex', *Geom. Funct. Anal.* **27**(6) (2017), 1450–1496.
[23] J.-O. Moussafir, 'On computing the entropy of braids', *Funct. Anal. Other Math.* **1**(1) (2006), 37–46.
[24] R. C. Penner, 'Probing mapping class groups using arcs', in *Problems on Mapping Class Groups and Related Topics*, Proc. Sympos. Pure Math. vol. 74 (Amer. Math. Soc., Providence, RI, 2006), 97–114.
[25] R. C. Penner and J. L. Harer, *Combinatorics of Train Tracks*, Annals of Mathematics Studies vol. 125 (Princeton University Press, Princeton, NJ, 1992).
[26] R. C. Penner, 'Bounds on least dilatations', *Proc. Amer. Math. Soc.* **113**(2) (1991), 443–450.
[27] R. C. Penner, *A Computation of the Action of the Mapping Class Group on Isotopy Classes of Curves and Arcs in Surfaces*, ProQuest LLC, Ann Arbor, MI, 1982. Thesis (Ph.D.)–Massachusetts Institute of Technology.
[28] K. Rafi, 'A characterization of short curves of a Teichmüller geodesic', *Geom. Topol.* **9** (2005), 179–202.
[29] R. Tang and R. C. H. Webb, 'Shadows of Teichmüller discs in the curve graph', *Int. Math. Res. Not. IMRN* **11** (2018), 3301–3341.
[30] W. P. Thurston, 'On the geometry and dynamics of diffeomorphisms of surfaces', *Bull. Amer. Math. Soc. (N.S.)* **19**(2) (1988), 417–431.
[31] W. P. Thurston, 'Three dimensional manifolds, Kleinian groups and hyperbolic geometry', in *Collected Works of William P. Thurston with Commentary . Vol. II. 3-manifolds, Complexity and Geometric Group Theory* (Amer. Math. Soc., Providence, RI, 2022), 9–33. Reprint of [0648524].
[32] Y. Wu, Personal communication, August 2024.
[33] C.-K. Yap et al., *Fundamental Problems of Algorithmic Algebra*, vol. 49 (Oxford University Press, Oxford, 2000).

https://doi.org/10.1017/fms.2026.10255 Published online by Cambridge University Press