The compression body graph has infinite diameter

Joseph Maher, Saul Schleimer

July 30, 2019

Abstract

We show that the compression body graph has infinite diameter.

Subject code: 37E30, 20F65, 57M50.

Contents

1 Introduction

The curve complex is a simplicial complex whose vertices are isotopy classes of simple closed curves, and whose simplices are spanned by simple closed curves which may be realized disjointly in the surface. In this paper, we consider a related complex, the *compression body graph*, defined by Biringer and Vlamis [BV17, page 94], which we shall denote by $\mathcal{H}(S)$. Vertices of this graph are isomorphism classes of *marked compression bodies.* A marked compression body (U, f) is a non-trivial compression body U together with a choice of homeomorphism $f: \partial_+ U \to S$, up to isotopy, from the upper boundary of the compression body to the surface S. Two marked compression bodies (U, f) and (V, g) are *isomomorphic* if they differ by a homeomorphism of the compression body. More precisely, there must be a homeomorphism $h: U \to V$ such that $f = g \circ h|_{\partial_+ U}$. Two compression bodies are *adjacent* if one is contained in the other one. More precisely, the compression body graph is a poset as follows: suppose that (U, f) and (V, g) are marked compression bodies, and $U' \subseteq U$ is a subcompression body of U. If there is a homeomorphism $h: U' \to V$ such that $f = g \circ h|_{\partial_+ U}$ then $V < U$. Biringer and Vlamis [BV17, Theorem 1.1], following Ivanov, showed that the simplicial automorphism group of this graph is equal to the mapping class group. We show:

Theorem 1.1. *The compression body graph $\mathcal{H}(S)$ is an infinite diameter Gromov hyperbolic metric space.*

The compression body graph $\mathcal{H}(S)$ is quasi-isometric to the metric space obtained by electrifying the curve complex along each *disc set*: the set of all simple closed curves that bound discs in a specific compression body. We will write $\pi_{\mathcal{Y}}$ for the inclusion map $X \hookrightarrow X_{\mathcal{Y}}$, which we shall also refer to as the *projection map.* As the electrification of a Gromov hyperbolic space along uniformly quasi-convex subsets is Gromov hyperbolic the compression body graph is Gromov hyperbolic. This follows from work of Bowditch [Bow12], Kapovich and Rafi [KR14] and Masur and Minsky [MM99, MM04]. See Section 2 for further details.

We also study the action of the mapping class group on $\mathcal{H}(S)$. In Lemma 7.2 we show that that for every genus $g \geqslant 2$, there are elements of the mapping class group of S which act loxodromically on $\mathcal{H}(S)$. Furthermore in Corollary 7.10 we show that every subgroup of the Johnson filtration contains elements which act loxodromically on $\mathcal{H}(S)$.

In the final section, we use our methods to give an alternate proof of Theorem 8.1: the stable lamination of a pseudo-Anosov element is contained in the limit set of a compression body V if and only if some power of the pseudo-Anosov extends over a non-trivial subcompression body of V. This was originally shown by Biringer, Johnson and Minsky [BJM13, Theorem 1], using techniques from hyperbolic three-manifolds. It has also been shown by Ackermann [Ack15, Theorem 1], extending the methods of Casson and Long [CL85]. We also weaken the hypotheses of a result of Lubotzky, Maher and Wu [LMW16, Theorem 1], showing that a random Heegaard splitting is hyperbolic with probability tending to one exponentially quickly, for a larger class of random walks than those considered in [LMW16].

The results of this paper have been used by Agol and Freedman [AF19], Burton and Purcell [BP14, Theorem 4.12], Dang and Purcell [DP17, Theorem 1.2] and Ma and Wang [MW17].

Finally, Theorem 1.1 implies that many graphs, formed by considering compression bodies of restricted topological type, are also of infinite diameter, see Section 6 for further details.

Acknowledgements

We thank Ian Biringer for enlightening conversations. The first author acknowledges support from the Simons Foundation and PSC-CUNY. The second author acknowledges support from the EPSRC (EP/I028870/1). The authors further acknowledge the support of the NSF (DMS-1440140) while the both were in residence at the MSRI in Berkeley, California, during the Fall 2016 semester.

2 Outline

We now give a brief outline of the main argument. The curve complex is Gromov hyperbolic, and the disc sets are quasi-convex. Thus, by electrifying the curve complex along the discs sets, we obtain a Gromov hyperbolic space $\mathcal{C}_{\mathcal{D}}(S)$ which is quasi-isometric to the compression body graph $\mathcal{H}(S)$. The electrification of a Gromov hyperbolic space along uniformly quasi-convex subsets is Gromov hyperbolic:

Theorem 2.1. *Let X be a Gromov hyperbolic space and let $\mathcal{Y} = \{Y_i\}_{i\in I}$ be a collection of uniformly quasi-convex subsets of X. Then $X_{\mathcal{Y}}$, the electrification of X with respect to $\mathcal{Y}$, is Gromov hyperbolic.*

Furthermore, for any constants K and c, there are constants K' and c' such that a (K,c)-quasi-geodesic in X projects to a reparameterized (K',c')-quasi-geodesic in $X_{\mathcal{Y}}$.

The first statement is due to Bowditch [Bow12, Proposition 7.12], and both statements follow immediately from Kapovich and Rafi [KR14, Corollary 2.4]. Masur and Minsky showed that the curve complex is Gromov hyperbolic [MM99, Theorem 1.1], and that the disc sets are quasi-convex [MM04, Theorem 1.1]. We shall write $\mathcal{C}(S)$ for the curve graph of the surface S, and $\mathcal{D}$ for the collection of all discs sets in $\mathcal{C}(S)$. Therefore $\mathcal{C}_{\mathcal{D}}(S)$ denotes the curve complex electrified along disc sets, which is quasi-isometric to the compression body graph $\mathcal{H}(S)$. These two quasi-isometric spaces are Gromov hyperbolic with infinite diameter by Theorem 1.1. Furthermore, work of Dowdall and Taylor [DT17, Theorem 3.2] shows that the Gromov boundary of $\mathcal{C}_{\mathcal{D}}(S)$ is homeomorphic to the subset of $\partial\mathcal{C}(S)$ consisting of the complement of the limit sets of disc sets.

Any quasi-geodesic ray $\gamma\colon \mathbb{N} \to \mathcal{C}(S)$ projects to a reparameterized and possibly finite diameter quasi-geodesic ray in the electrification $\mathcal{C}_{\mathcal{D}}(S)$. In Section 4, we prove our stability result for quasi-geodesics in the curve complex $\mathcal{C}(S)$: if $\pi_{\mathcal{H}} \circ \gamma$ has finite diameter image in $\mathcal{C}_{\mathcal{D}}(S)$ then there is a constant k, and a handlebody V, such that γ is contained in a k-neighbourhood of the disc set $\mathcal{D}(V)$ in $\mathcal{C}(S)$. In particular, this means that the ending lamination corresponding to γ lies in the limit set of the disc set $\mathcal{D}(V)$. However, for a given handlebody V, the limit set of its disc set $\mathcal{D}(V)$ in the boundary of the curve complex has measure zero. As there are only countably many discs sets, the union of their limit sets also has measure zero. Therefore, there is a full measure set of minimal laminations disjoint from the union of the disc sets. Any one of these gives rise to a geodesic ray whose image in the electrification $\mathcal{C}_{\mathcal{D}}(S)$ has infinite diameter. In particular, $\mathcal{C}_{\mathcal{D}}(S)$ has infinite diameter.

We now give a brief sketch of the proof of the stability result. Suppose there is a geodesic ray γ in the curve complex $\mathcal{C}(S)$, and a sequence of handlebodies V_i, such that the initial segment of γ of length i is contained in a k'-neighbourhood of $\mathcal{D}(V_i)$. In particular, for any i, there are infinitely many disc sets $\mathcal{D}(V_j)$ passing within distance k' of both γ_0 and γ_i.

Recall that given two simple closed curves a and b on a surface S, we may *surger* a, along an innermost arc of b, to produce a new curve a', disjoint from a, and with smaller geometric intersection number with b. This is illustrated in Figure 6. By iterating this procedure, we obtain a surgery sequence, which gives rise to a reparameterized quasi-geodesic in $\mathcal{C}(S)$ from a to b. We call this a *curve surgery sequence*. If the simple closed curves a and b bound discs in a handlebody V, then we may surger a along innermost bigons of b, to produce a surgery sequence connecting a and b, in which every surgery curve bounds a disc in V. We shall call such a surgery sequence a *disc surgery sequence*.

Recall that a *train track* on a surface S is a smoothly embedded graph, such that the edges at each vertex are all mutually tangent, and there is at least one edge in each of the two possible directed tangent directions. Furthermore, there are no complementary regions which are nullgons, monogons, bigons or annuli.

A *split* of a train track τ is a new train track τ' obtained by one of the local modifications illustrated in Figure 2. We say that a sequence of train tracks $\{\tau_j\}$ is a *splitting sequence* if each τ_{j+1} is obtained as a split of τ_j.

A key result we use from [MM04, page 319] says, roughly speaking, that we may choose a surgery sequence $\{D_j\}$ connecting two discs in a common compression body such that there is a corresponding train track splitting sequence $\{\tau_j\}$, so that for each j, the disc D_j is dual to the train track τ_j, and meets it in a single point at the switch. We state a precise version of this as Proposition 3.13 below.

Recall that given an essential subsurface X contained in S, there is a *subsurface projection map* from a subset of $\mathcal{C}(S)$ to $\mathcal{C}(X)$. Roughly speaking, this map sends a simple closed curve a, meeting X essentially, to a simple closed curve $a' \subset X$, which is disjoint from some arc of $a \cap X$. See Section 3 for a precise definition. We may extend the definition of subsurface projection from simple closed curves to train tracks, by instead projecting the *vertex cycles* of the train track τ. See Section 3.4 for further details. The collection of vertex cycles $\{\Lambda(\tau_j)\}$ for a splitting sequence $\{\tau_j\}$ projects to a reparameterized quasi-geodesic in $\mathcal{C}(S)$.

We say that three (ordered) points x, y and z in a metric space satisfy the *reverse triangle inequality* with constant K if $d(x, y) + d(y, z) \leqslant d(x, z) + K$. By the Morse lemma, given constants δ, Q and c, there is a constant K, such that if y lies on a (Q, c)-quasi-geodesic between x and y in a δ-hyperbolic space, then x, y and z satisfy the reverse triangle inequality. Given three (ordered) points x, y and z, we say that y is *K-intermediate* with respect to x and z, if x, y and z satisfy the reverse triangle inequality, and furthermore $d(x, y) \geqslant K$ and $d(y, z) \geqslant K$.

Given a train track sequence $\{\tau_j\}$, we say that a particular train track τ_j is *K-intermediate* if the vertex cycles of τ_j are distance at least K in the curve complex from the vertex cycles of both τ_0 and τ_n. If K is sufficiently large, then for any two train track sequences $\{\tau_j\}$ and $\{\tau'_j\}$ starting near γ_0 and ending near γ_r, for any subsurface X in S, and for any pair of K-intermediate train tracks τ_j and τ'_k, the distance between the subsurface projections of τ_j and τ'_k in $\mathcal{C}(X)$ is bounded in terms of $d_X(\gamma_0, \gamma_r)$. In particular, using the Masur-Minsky distance formula we find that every train track sequence connecting γ_0 and γ_r passes within a bounded marking distance of any K-intermediate train track. As the marking complex is locally finite, infinitely many train track sequences must share a common train track, and this implies that infinitely many of the handlebodies V_i must share a common disc D_r. Furthermore, the boundary of D_r lies close to the geodesic from γ_0 to γ_r.

We iterate this argument to obtain the following. Choose an increasing sequence of numbers $\{r_n\}$, with the difference between consecutive numbers fixed and sufficiently large. At each stage, we may assume we have passed to an infinite subset of the handlebodies so that each contains the discs $D_{r_1}, \ldots D_{r_n}$. These discs lie in a common compression body W_n, and the geodesic from γ_0 to γ_{r_n} is contained in a bounded neighbourhood of $\mathcal{D}(W_n)$. The compression bodies W_n form an increasing sequence $W_n < W_{n+1}$, which eventually stabilizes to a constant sequence W. The entire geodesic ray γ is then contained in a bounded neighbourhood of $\mathcal{D}(W)$, by the stability result.

In the next section, Section 3, we review the results we will use, and set up some notation. In Section 4, we provide the details of the proof of the stability result. In Section 5, we use the stability result to prove Theorem 1.1. In Section 7, we show that there are many loxodromic isometries of the compression body graph. In the final section, Section 8, we give several applications.

3 Preliminaries

In this section we review some background material, and set up some notation.

3.1 The mapping class group

Let S be a compact connected oriented surface, possibly with boundary. The *mapping class group* $\mathrm{Mod}(S)$ is the group of homeomorphisms of S, up to isotopy. We shall fix a finite generating set for the mapping class group, and we shall write $d_{\mathrm{Mod}}(g, h)$ for the corresponding word metric on $\mathrm{Mod}(S)$.

We say a simple closed curve in S is *peripheral* if it cobounds an annulus together with one of the boundary components of S. We say a simple closed curve in S is *essential* if it does not bound a disc and is not peripheral.

Given two finite collections of essential curves μ and μ', we may extend the geometric intersection number from single curves to finite collections by

$$i(\mu, \mu') = \sum_{a \in \mu,\, a' \in \mu'} i(a, a').$$

We say a set of curves μ is a *marking* of S if the set of curves *fills* the surface: that is, for all curves $a \in \mathcal{C}(S)$ we have $i(a, \mu) > 0$. We say a marking μ is an L-*marking* if $i(\mu, \mu) \leqslant L$. We say two markings μ and μ' are L'-adjacent if $i(\mu, \mu') \leqslant L'$. Define $\mathcal{M}_{L,L'}(S)$ to be the graph whose vertices are (isotopy classes of) L-markings and whose edges connect pairs of L-markings which are L'-adjacent. Masur and Minsky [MM00, Section 7.1] show that, for sufficiently large constants L and L', the *marking graph* $\mathcal{M}(S) = \mathcal{M}_{L,L'}(S)$ is locally finite, connected and quasi-isometric to the Cayley graph of the mapping class group $(\mathrm{Mod}(S), d_{\mathrm{Mod}})$. We shall fix a pair of constants L and L' with this property for the remainder of this paper, and will suppress these constants from our notation.

The *complex of curves* $\mathcal{C}(S)$ is a simplicial complex, whose vertices consist of isotopy classes of essential curves, and whose simplices are spanned by disjoint (perhaps after isotopy) curves. We need to modify this definition for certain low-complexity surfaces, as we now describe. In the case of a once-holed torus we connect two curves by an edge if they have geometric intersection number one. In the case of a four-holed sphere we connect two curves by an edge if they have geometric intersection number two. Finally, if the surface S is an annulus A, then we define the curve complex $\mathcal{C}(A)$ as follows: the vertices consist of properly embedded essential arcs up to homotopies fixing the endpoints, and two arcs are connected by an edge if they may be realized disjointly in the interior of the annulus. We shall consider the complex of curves as a metric space in which each edge has length one. We will write d_S for distance in the complex of curves. We extend the definition of the metric from points to finite sets by setting

$$d_S(A, B) = \min_{a \in A,\, b \in B} d_S(a, b).$$

Suppose that S is not an annulus. Then a subsurface $Y \subset S$ is *peripheral* if it is an annulus with peripheral boundary. A subsurface $Y \subset S$ is *essential* if it is not peripheral and if all boundary components are essential or peripheral. Let $Y \subset S$ be an essential subsurface. We will write d_Y for the metric in $\mathcal{C}(Y)$. Let $Y_\varnothing$ be the set of vertices of $\mathcal{C}(S)$ corresponding to essential curves which may be isotoped to be disjoint from Y. There is a coarsely well defined *subsurface projection* $\pi_Y \colon \mathcal{C}(S) - Y_\varnothing \to \mathcal{C}(Y)$, which we now define. We say a properly embedded arc in a surface S is *essential* if it does not bound a properly embedded bigon together with a subarc of the boundary of S. Let $\mathcal{AC}(S)$ be the *arc and curve complex* of S: the simplicial complex whose vertices are isotopy classes of essential curves and properly embedded essential simple arcs. Thus $\mathcal{AC}(S)$ contains $\mathcal{C}(S)$ as a subcomplex, and this inclusion is a quasi-isometry. Let S^Y be the cover of S corresponding to Y. The surface Y is homeomorphic to the Gromov closure of S^Y so we may identify $\mathcal{AC}(Y)$ with $\mathcal{AC}(S^Y)$. For any essential curve or arc a, let a^Y be the full preimage of a in S^Y. Define $\pi_Y(a)$ to be the set of essential components of a^Y in S^Y. So $\pi_Y(a)$ is either empty, or is a simplex in $\mathcal{AC}(Y)$. As $\mathcal{AC}(Y)$ is quasi-isometric to $\mathcal{C}(Y)$, this gives a coarsely well-defined map to $\mathcal{C}(Y)$ in the latter case.

We define the *cut-off function* $\lfloor \cdot \rfloor_c$ on $\mathbb{R}$ by $\lfloor x \rfloor_c = x$ if $x \geqslant c$ and zero otherwise. Given a set X, two functions f and g on $X \times X$, and two constants $K \geqslant 1$ and $c > 0$, we say that f and g are (K,c)-coarsely equivalent, denoted $f \approx_{(K,c)} g$, if for all x and y in X we have

$$\frac{1}{K} f(x,y) - \frac{c}{K} \leqslant g(x,y) \leqslant K f(x,y) + c.$$

We now state the *distance estimate* due to Masur and Minsky.

Theorem 3.1. [MM00, Theorem 6.12] *For any surface S there is a constant M_0, such that for any $M \geqslant M_0$, there are constants K and c, such that for any markings μ and ν,*

$$d_{\mathcal{M}(S)}(\mu,\nu) \approx_{(K,c)} \sum_X \lfloor d_X(\mu,\nu) \rfloor_M.$$

That is, the distance in the marking complex is coarsely equivalent to the (cut-off) sum of subsurface projections. Note that there are only finitely many non-zero terms on the right-hand side.

3.2 Compression bodies and discs sets

Let S be a closed connected oriented surface of genus g. A *compression body* V is a compact orientable three-manifold obtained from $S \times I$ by attaching two-handles to $S \times \{0\}$, and capping off any newly created two-sphere components with three-balls. In particular, if S is a two-sphere, we do not cap off $S \times \{1\}$. The genus g surface $S \times \{1\}$ is the *upper boundary* $\partial_+ V$, while the other boundary components make up the *lower boundary* $\partial_- V$. The lower boundary need not be connected. If $\partial_- V$ is empty, then V is a *handlebody.* The *trivial compression body* has no two-handles, and is homeomorphic to $S \times I$.

A *marked compression body* is a pair (V, f) where V is a compression body and $f \colon \partial_+ V \to S$ is a homeomorphism. Two marked compression bodies (V, f) and (W, g) are *isomorphic* if there is a homeomorphism $h \colon V \to W$ such that $g \circ (h|_{\partial V}) = f$. In what follows, we will suppress the marking f and assume that ∂V and S are actually equal.

Biringer and Vlamis define the *compression body graph* $\mathcal{H}(S)$ to be the graph whose vertices are isomorphism classes of non-trivial marked compression bodies. Here V and W are adjacent if either $V < W$ or $W < V$ [BV17, page 94]. They show that the simplicial automorphism group of $\mathcal{H}(S)$ is equal to the mapping class group $\mathrm{Mod}(S)$ [BV17, Theorem 1.1] for genus at least three. In genus two, the mapping class group surjects onto the automorphism group, with kernel generated by the hyperelliptic involution.

We say a disc D in a marked compression body (V, f) is *essential* if it is properly embedded in V. In this case, its boundary is an essential curve in the upper boundary $\partial_+ V$. We say an essential curve in S bounds a disc in V if it is the image under f of the boundary of an essential disc in V.

Definition 3.2. The *disc set* $\mathcal{D}(V)$ of a marked compression body (V, f) is the collection of all essential curves in S which bound discs in V.

A subset Y of a geodesic metric space is *Q-quasi-convex* if for any pair of points x and y in Y, any geodesic connecting x and y is contained in a Q-neighbourhood of Y. Masur and Minsky [MM04, Theorem 1.1] showed that there is a Q such that the disc set $\mathcal{D}(V)$ is a Q-quasi-convex subset of $\mathcal{C}(S)$. Given a metric space (X, d), and a collection of subsets $Y = \{Y_i\}_{i \in I}$, the *electrification* of X with respect to Y is the metric space $X_{\mathcal{Y}}$ obtained by adding a new vertex y_i for each set Y_i, and coning off Y_i by attaching edges of length $\frac{1}{2}$ from each $y \in Y_i$ to y_i; the image of each set Y_i in $X_{\mathcal{Y}}$ has diameter one. We shall write $\mathcal{D}(S)$ to denote the collection of all disc sets of all non-trivial compression bodies V with boundary S. The compression body graph $\mathcal{H}(S)$ is quasi-isometric to the curve complex electrified along all the disc sets, namely $\mathcal{C}(S)_{\mathcal{D}(S)}$.

We remark that there is another natural space quasi-isometric to $\mathcal{C}(S)_{\mathcal{D}(S)}$, known as the *graph of handlebodies*, which has vertices being classes of marked handlebodies and edges being distinct pairs $\{V, W\}$ where $\mathcal{D}(V)$ has non-empty intersection with $\mathcal{D}(W)$.

3.3 Surgery sequences

In this section we recall the definition of surgery sequences for simple closed curves and for discs. Since we restrict our attention to *closed* surfaces our discussion is simpler than the more general case of compact surfaces with boundary. In subsequent sections we may abuse notation by referring to these as just surgery sequences, if it is clear from context whether we mean curves or discs.

Definition 3.3. Let a be an essential curve in S, and let b' be a simple arc whose endpoints lie on a, and whose interior is disjoint from a. Furthermore, suppose that b' is essential in $S - a$. The endpoints of b' divide a into two arcs with common endpoints, a' and a'', say. A simple closed curve c is said to be produced by *(arc) surgery of a along b'* if c is homotopic to either of the simple closed curves $a' \cup b'$ or $a'' \cup b'$.

Definition 3.4. Let a and b be essential simple closed curves in minimal position with $i(a, b) \geqslant 2$. An *innermost arc* of b with respect to a is a subarc $b' \subset b$ whose endpoints lie on a, and whose interior is disjoint from a. A simple closed curve c is said to be produced by *(curve) surgery of a along b* if c is produced by surgery of a along b', for some choice of innermost arc b' in b.

If c is produced by surgery of a along b, then the number of intersections of c with b is strictly less than the number of intersections of a with b.

Definition 3.5. Given a pair of essential simple closed curves a and b, a *(curve) surgery sequence* connecting a and b is a sequence of simple closed curves $\{a_i\}_{i=0}^{n-1}$, such that $a_0 = a$, the final curve a_{n-1} is disjoint from b, and each a_{i+1} is produced from a_i by a surgery of a_i along b.

Definition 3.6. Let D and E be essential discs in a compression body V in minimal position, and let E' be an outermost bigon of E with respect to the arcs of $D \cap E$. The arc of intersection between E' and D divides D into two discs, D' and D'', say. We say that a disc F is produced by *(disc) surgery of D along E'* if F is homotopic to either of the discs $D' \cup E'$ or $D'' \cup E'$, for some choice of outermost bigon E' contained in E.

The disc F produced by surgery of D along E' is disjoint from D, and is essential. To see this, suppose that one of the discs F is inessential. That is, ∂F bounds a disc C in S. We can then ambiently isotope $\partial E' \cap S$ across C, reducing the number of intersections between ∂D and ∂E, a contradiction.

Definition 3.7. Given a pair of essential discs D and E in minimal position, a *(disc) surgery sequence* connecting D and E is a sequence of essential properly embedded discs $\{D_i\}_{i=0}^{n-1}$, such that $D_1 = D$, the final disc D_{n-1} is disjoint from E, and each D_{i+1} is produced from D_i by a (disc) surgery of D_i along E.

We remark that if $\{D_i\}_{i=1}^{n}$ is a (disc) surgery sequence, then $\{\partial D_i\}_{i=1}^{n}$ is a (curve) surgery sequence.

Proposition 3.8. *Let V be a compression body. Any two curves in its disc set $\mathcal{D}(V)$ are connected by a disc surgery sequence.*

Proof. Let D and E be two essential discs in a compression body V, and assume they intersect in minimal position. Choose an outermost bigon E' of E with respect to the arcs $D \cap E$. We may surger the disc D along E', which produces two new discs in V disjoint from D. Call one of these D_1. The disc D_1 is disjoint from D and has fewer intersections with E. This process terminates after finitely many disc surgeries. This gives a disc surgery sequence $\{D_i\}_{i=0}^{n-1}$ connecting D and E. $\square$

3.4 Train tracks

We briefly recall some of the results we use about train tracks on surfaces. For more details see for example Penner and Harer [PH92].

A *pre-train track* τ is a smoothly embedded finite graph in S such that the edges at each vertex are all mutually tangent, and there is at least one edge in each of the two possible directed tangent directions. The vertices are commonly referred to as *switches* and the edges as *branches*. We will always assume that all switches have valence at least three. A trivalent switch is illustrated below in Figure 1. If none of the complementary regions of τ in S are nullgons, monogons, bigons or annuli, then we say that τ is a *train track*. Up to the action of the mapping class group, there are only finitely many train tracks in S.

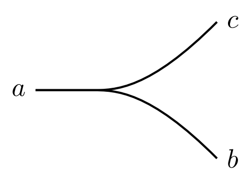


Figure 1: A trivalent switch for a train track.

An assignment of non-negative numbers to the branches of τ, known as *weights*, satisfies the *switch equality* if the sum of weights in each of the two possible directed tangent directions is equal: that is $a = b + c$ in Figure 1 above. A weighted train track defines a measured lamination on the surface. We say that the corresponding lamination is *carried* by the train track. A train track τ determines a polytope of projectively measured laminations $P(\tau) \subset \mathcal{PML}(S)$ carried by τ. Let $\Lambda(\tau)$ be the set of vertices of $P(\tau)$. Every $v \in \Lambda(\tau)$ gives a *vertex cycle*: a simple closed curve, carried by τ, that puts weight at most two on each branch of τ. It follows that the set $\Lambda(\tau)$ gives a finite set of curves in $\mathcal{C}(S)$ and these curves have bounded pairwise geometric intersection numbers. A train track τ is *filling* if $\Lambda(\tau)$ is a marking. As there are only finitely many train tracks up to the action of the mapping class group, these bounds depend only on the surface S.

A simple closed curve a is *dual* to a track τ if a misses the switches of τ, crosses the branches of τ transversely, and forms no bigons with τ. We say that a simple closed curve a is *switch-dual* to τ if a meets the train track transversely at exactly one point of τ, which is a switch, and forms no bigons with τ.

A track τ is *large* if all components of $S - \tau$ are discs or peripheral annuli. A track τ is *recurrent* if for every branch $b \subset \tau$ there is a curve $\alpha \in P(\tau)$ putting positive weight on b. A track τ is *transversely recurrent* if for every branch $b \subset \tau$ there is a curve β dual to τ, such that β crosses b.

Lemma 3.9. *Given a surface S, there is a constant N, with the following property. For any marking μ of S, there are at most N non-isotopic filling train tracks τ with $\Lambda(\tau) = \mu$.*

Proof. There are only finitely many isotopy classes of filling train tracks up to the action of the mapping class group. If $\tau = g\sigma$, for some element of the mapping class group $g \in \mathrm{Mod}(S)$, and $\Lambda(\tau) = \Lambda(\sigma) = \mu$, then g preserves the collection of curves μ. Note that the collection of curves μ fills S and has finite total self-intersection number depending only on S. We deduce that there are only finitely many such mapping classes g. $\square$

A *split* of a train track τ produces a new train track σ by one of the local modifications illustrated in Figure 2 below. Here a subset of τ diffeomorphic to the top configuration, is replaced with one of the lower three configurations.

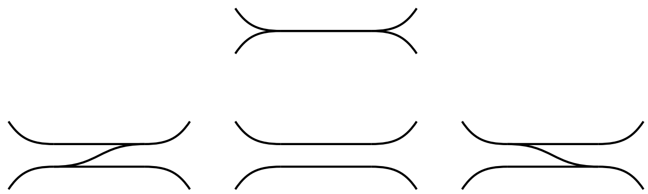


Figure 2: Splitting a train track.

A *shift* for a train track is the local modification given in Figure 3.

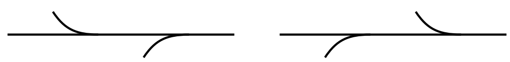


Figure 3: Shifting for a train track.

A *tie neighbourhood* $N(\tau)$ for a train track τ is a union of rectangles, as follows. For each switch there is a rectangle $R(s)$, and for each branch b there is a rectangle $R(b)$. All rectangles are foliated by vertical arcs called *ties*. We glue a vertical side of a branch rectangle to a subset of a vertical side of a switch rectangle as determined by the combinatorics of τ. See Figure 4 for a local picture near a trivalent switch.

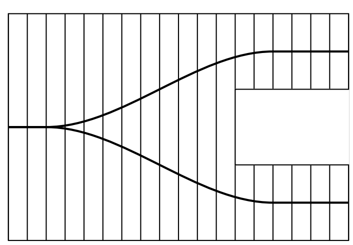


Figure 4: A tie neighbourhood for a train track.

We say a train track or simple closed curve σ is *carried* by τ if σ may be isotoped to lie in $N(\tau)$, such that σ is transverse to the ties of $N(\tau)$. We denote this by either $\sigma \prec \tau$ or $\tau \succ \sigma$. If σ is a train track obtained by splitting and shifting τ then $\tau \succ \sigma$. A *train track carrying sequence* is a sequence of train tracks $\tau_0 \succ \tau_1 \succ \cdots$, such that each τ_i carries τ_{i+1}. We may also denote a train track splitting sequence by $\{\tau_i\}_{i=0}^n$, where $n \in \mathbb{N}_0 \cup \{\infty\}$. We say that $\{\tau_i\}_{i=0}^n$ is *K-connected* if

$$d_S(\Lambda(\tau_i), \Lambda(\tau_{i+1})) \leqslant K$$

for all i.

Masur and Minsky [MM04, Theorem 1.3] showed that train track splitting sequences give rise to reparameterized quasi-geodesics in the curve complex

$\mathcal{C}(S)$. Using this, Masur, Mosher and Schleimer show that train track splitting sequences give reparameterized quasi-geodesics under subsurface projection.

Theorem 3.10. [MMS12, Theorem 5.5] *For any surface S with $\xi(S) \geqslant 1$ there is a constant $Q = Q(S)$ with the following property: For any sliding and splitting sequence $\{\tau_i\}_{i=0}^n$ of birecurrent train tracks in S and for any essential subsurface $X \subset S$, if $\pi_X(\tau_i) \neq \varnothing$ then the sequence $\{\pi_X(\Lambda(\tau_i))\}_{i=0}^n$ is a Q-reparameterized quasi-geodesic in the curve complex $\mathcal{C}(X)$.*

We abuse notation by writing $d_X(\tau, \cdot)$ for $d_X(\pi_X(\Lambda(\tau)), \cdot)$, where X is an essential subsurface of S. Given a train track carrying sequence $\{\tau_i\}_{i=0}^n$, we say that a train track τ_i in the sequence is *K-intermediate* if $d_S(\tau_0, \tau_i) \geqslant K$ and $d_S(\tau_i, \tau_n) \geqslant K$. We will use the following consequence of Theorem 3.10:

Theorem 3.11. *There is a constant B, depending only on S, such that for any constant A, and any two carrying sequences $\{\tau_i\}_{i=0}^n$ and $\{\sigma_i\}_{i=0}^m$, with $d_S(\tau_0, \sigma_0) \leqslant A$ and $d_S(\tau_n, \sigma_m) \leqslant A$, then for any pair of $(A + B)$-intermediate train tracks τ_i and σ_j, and any subsurface $X \subset S$,*

$$d_X(\tau_i, \sigma_j) \leqslant d_X(\tau_0, \tau_n) + B.$$

To prove Theorem 3.11 we will also need the following *bounded geodesic image theorem* of Masur and Minsky.

Theorem 3.12. [MM00, Theorem 3.1] *For any surface S, and for any essential subsurface X of S, there is a constant M_X, such that for any geodesic γ in $\mathcal{C}(S)$, all of whose vertices intersect X non-trivially, the projected image of γ in $\mathcal{C}(X)$ has diameter at most M_X.*

Proof of Theorem 3.11. If there is a constant B, depending on S, such that $d_X(\tau_i, \sigma_j) \leqslant B$, then we are done. Otherwise, by the bounded geodesic image theorem, ∂X is also $(A + B)$-intermediate. Then $d_X(\tau_0, \sigma_0) \leqslant M_X$ and $d_X(\tau_n, \sigma_m) \leqslant M_X$.

By Theorem 3.10, there is a constant Q, such that the images of both $\{\Lambda(\tau_i)\}_{i=1}^n$ and $\{\Lambda(\sigma_i)\}_{i=1}^m$ under π_X are reparameterized Q-quasi-geodesics, and furthermore, their endpoints are distance at most A_2 apart in $\mathcal{C}(X)$. By the Morse property for quasi-geodesics, there is a constant A_3, depending only on M_X, Q and δ, and hence only on the surface S, such that $\{\Lambda(\tau_i)\}_{i=1}^n$ and $\{\Lambda(\sigma_i)\})_{i=1}^m$ are Hausdorff distance at most A_3 apart, and so the distance between $\pi_X(\tau_i)$ and $\pi_X(\sigma_j)$ is at most $d_X(\tau_0, \tau_n) + A_3$. Therefore the result follows, choosing $B = \max\{A_1 + O(\delta), A_3\}$, which only depends on the topology of the surface S. □

We will abuse notation and say that an essential disc E is *carried* by a train track τ if ∂E is carried by τ.

We will use the following result of Masur and Minsky [MM04, page 309].

Proposition 3.13. *Let D and E be essential discs contained in a compression body V. Then there is a surgery sequence $\{D_i\}_{i=0}^{n-1}$ with $D = D_0$ and D_{n-1} disjoint from E, and a carrying sequence $\{\tau_i\}_{i=0}^n$, with the following properties.*

1. *The train track τ_i has only one switch for all $0 \leqslant i \leqslant n$.*

2. *The boundary of D_i is switch-dual to the track τ_i, for all $0 \leqslant i < n$.*

3. *The disc E is carried by the train track τ_i, for all $0 \leqslant i \leqslant n$.*

Masur and Minsky [MM04] work in a more general setting allowing for surfaces with boundary components, and state a weaker version of property 2, that $d_S(D_i, \tau_i) \leqslant 5$. However, in the case of closed surfaces, the statement we need follows from the proof of [MM04, Lemma 4.1]. We provide a review of their work in the appendix for the convenience of the reader.

Finally, we observe:

Proposition 3.14. *For all surfaces S, there is a constant K such that for any curves ∂D and ∂E in $\mathcal{D}$, and any carrying sequence $\{\tau_i\}_{i=0}^n$ connecting them, any K-intermediate train track in $\{\tau_i\}_{i=0}^n$ is birecurrent and filling.*

Proof. For all i, the curve ∂E runs over every edge of τ_i, so τ_i is recurrent. The train track τ_i has only one switch, and ∂D_i crosses τ_i exactly once at that switch, so τ_i is transversely recurrent.

If τ_i is not filling, then there is a curve a disjoint from the set of vertex cycles $\Lambda(\tau_i)$. Then a is disjoint from every collection of simple closed curves produced from positive integer valued sums of the $\Lambda(\tau_i)$, and in particular τ_i cannot carry a pair of curves which fill S. However, the train track τ_i carries ∂E. As long as $d_S(\Lambda(\tau_i), \partial E) \geqslant 3$, there is a vertex cycle distance at least 3 from ∂E, and so τ_i carries a pair of curves which fill S, and so $\Lambda(\tau_i)$ fills. Therefore a K-intermediate train track τ_i is filling, for any $K \geqslant 3$. □

3.5 Laminations

If f is a pseudo-Anosov homeomorphism of S which extends over a compression body V, then the stable lamination of f is a limit of discs of V. On the other hand we have the following:

Proposition 3.15. *There is a minimal lamination in $\mathcal{PML}(S)$ which does not lie in the limit set of a disc set $\mathcal{D}(V)$, for any compression body V.*

Proof. Recall that $\mathcal{PML}(S)$ is the space of projectively measured laminations in S. Fix a handlebody V. Define the *limit set* $\overline{\mathcal{D}}(V)$ of V to be the closure of $\mathcal{D}(V)$, considered as a subset of $\mathcal{PML}(S)$. Following Masur [Mas86, Theorem 1.2], we define $Z(V)$, the *zero set* of $\overline{\mathcal{D}}(V)$, to be the set of laminations $\mu \in \mathcal{PML}(S)$ where there is some $\nu \in \overline{\mathcal{D}}(V)$ having geometric intersection $i(\mu, \nu)$ equal to zero. Note that $Z(V)$ is closed in $\mathcal{PML}(S)$. Masur [Mas86, Theorem 1.2] shows that $Z(V)$ has empty interior in $\mathcal{PML}(S)$, and Gadre [LM12, Theorem A.1] shows that it has measure zero. Therefore, the complement of the countable union $\cup_V Z(V)$ is full measure in $\mathcal{PML}(S)$. The set of minimal laminations also has full measure in $\mathcal{PML}(S)$, so there is at least one minimal lamination disjoint from the union of the limit sets of the disc sets. □

We remark that this argument also works using harmonic measure instead of Lebesgue measure by work of Kaimanovich and Masur [KM96, Theorem 2.2.4] and Maher [Mah10, Theorem 3.1].

4 Stability

Theorem 4.1. *Suppose $\gamma\colon \mathbb{N} \to \mathcal{C}(S)$ is a geodesic ray. The diameter of the image of $\pi_{\mathcal{H}(S)} \circ \gamma$ is finite if and only if there is a compression body V and a constant k so that the image of γ lies in a k-neighborhood of $\mathcal{D}(V)$.*

The backward direction is immediate. In this section we show the forward direction. A standard coarse geometry argument using the hyperbolicity of the curve complex, plus the quasi-convexity of $\mathcal{D}(V)$ inside of $\mathcal{C}(S)$, reduces the forward direction of Theorem 4.1 to the following statement, which we call the stability hypothesis.

Theorem 4.2 (Stability hypothesis)**.** *Given a surface S and a constant k, there is a constant $k' \geqslant k$ with the following property. Suppose that γ is a geodesic ray in $\mathcal{C}(S)$, and V_i is a sequence of compression bodies such that, for all i, the segment $\gamma|[0,i]$ lies in a k–neighborhood of $\mathcal{D}(V_i)$. Then there is a constant k' and a non-trivial compression body W, contained in infinitely many of the V_i, such that that γ is contained in a k'-neighborhood of $\mathcal{D}(W)$.*

We first show the following.

Lemma 4.3. *There is a constant K, which depends on S, such that for any two essential simple closed curves a and b in S, with $d_S(a,b) \geqslant 3K$, there is a constant $N(a,b,K)$, such that for any collection $\{V_i\}_{i=1}^{N}$ of compression bodies with $d_S(a,\mathcal{D}(V_i)) \leqslant K$ and $d_S(b,\mathcal{D}(V_i)) \leqslant K$ there are at least two compression bodies V_i and V_j which share a common simple closed curve c. Furthermore, the curve c is K-intermediate for a and b.*

For fixed K, the constant N is coarsely equivalent to the smallest marking distance between any markings containing a and b.

Proof of Lemmma 4.3. For each compression body V_i, choose curves a_i and b_i in $\mathcal{D}(V)$ such that $d_S(a,a_i) \leqslant K$ and $d_S(b,b_i) \leqslant K$. The discs bounded by a_i and b_i in V_i determine a disc surgery sequence $\{D^i_j\}_{j=0}^{n_i-1}$, with $D^i_0 = a_i$ and $D^i_{n_i-1}$ disjoint from b_i, and a train track carrying sequence $\{\tau^i_j\}_{j=0}^{n_i-1}$.

By Proposition 3.14 there is a K such that for all i,j, every K-intermediate train track τ^i_j is filling and birecurrent. By Theorem 3.11, and bounded geodesic projections, there is a constant K_1 such that for all K-intermediate train tracks $\tau^{i_1}_{j_1}$ and $\tau^{i_2}_{j_2}$, and any subsurface $Y \subset S$,

$$d_Y(\tau^{i_1}_{j_1}, \tau^{i_2}_{j_2}) \leqslant d_Y(a,b) + K_1.$$

Therefore, using the Masur-Minsky distance formula, Theorem 3.1, with cutoff M larger than $K_1 + M_0$, there is a constant N_1 such that

$$d_{\mathcal{M}}(\tau_{j_1}^{i_1}, \tau_{j_2}^{i_2}) \leqslant N_1.$$

By Lemma 3.9 there are most N_2 train tracks τ with $\Lambda(\tau) = \mu$, for any marking μ. As $d_{\mathcal{M}}$ is a proper metric, for any K-intermediate track τ_j^i, there are at most $N_3 = N_2 |B_{\mathcal{M}}(\Lambda(\tau_j^i), N_1)|$ train tracks with markings within distance N_1 of $\Lambda(\tau_j^i)$. Thus, if there are at least $N_3 + 1$ compression bodies, at least two of them must share a common train track.

For each train track τ_j^i, there is a unique simple closed curve that is switch-dual to τ_j^i, and bounds a disc in the compression body V_i. Therefore, if there are at least $N = N_3 + 1$ compression bodies, at least two of them must have a disc in common. □

We now complete the proof of Theorem 4.2.

Proof of Theorem 4.2. Choose a subsequence $(n_k)_{k \in \mathbb{N}_0}$ with $n_0 = 0$ such that $d_S(\gamma(n_k), \gamma(n_{k+1})) \geqslant 3K$ for all k.

By Lemma 4.3, there is an infinite subset of the compression bodies V_i such that the V_i contain a simple closed curve ∂D_0 which is K-intermediate for $\gamma(n_0)$ and $\gamma(n_1)$. We may pass to this infinite subset, and then apply Lemma 4.3 to $\gamma(n_1)$ and $\gamma(n_2)$, producing a simple closed curve ∂D_1, which is K-intermediate for $\gamma(n_1)$ and $\gamma(n_2)$, and such that ∂D_0 and ∂D_1 simultaneously compress in infinitely many compression bodies.

Isotope ∂D_0 and ∂D_1 in S so they realize their geometric intersection number. By work of Casson and Long [CL85, Proof of Lemma 2.2], the curves ∂D_0 and ∂D_1 simultaneously compress in a handlebody V if and only if there is a pairing on the points of $\partial D_0 \cap \partial D_1$ that is simultaneously unlinked on ∂D_0 and on ∂D_1. By the previous paragraph, we know that there is at least one such pairing. There are only finitely many such pairings, so we may pass to a further subsequence of the V_i where ∂D_0 and ∂D_1 compress in all of the V_i, and with the same pairing. It follows that these V_i share a common non-trivial compression body W_1, containing ∂D_0 and ∂D_1.

We now iterate the argument, finding simultaneous compressions $\partial D_1, \partial D_2, \partial D_3, \ldots$ for a descending chain of subsequences of $\{V_i\}$. This gives rise to an ascending chain of compression bodies $W_1 \subset W_2 \subset W_3 \subset \ldots$. However, an ascending chain of compression bodies must stabilize after finitely many steps.

If there is some m so that W_m is a handlebody, then infinitely many of the V_i are equivalent. If there is some m so that $W_m = W_n$ for all $n > m$ then the compression body $W = W_m$ is contained in infinitely many of the V_i. In either case we have completed the proof of Theorem 4.2. □

5 Infinite diameter

We will use the following result of Klarreich [Kla, Theorem 1.3]. See also Hamenstädt [Ham06].

Theorem 5.1. [Kla, Theorem 1.3] *The Gromov boundary of the complex of curves $\mathcal{C}(S)$ is homeomorphic to the space of minimal foliations on S.*

We may now complete the proof of Theorem 1.1.

Proof of Theorem 1.1. Fix a geodesic ray $\gamma\colon \mathbb{N} \to \mathcal{C}(S)$, and set $\gamma_i = \gamma(i)$. Applying Theorem 5.1, let λ be the ending lamination associated to γ. That is, after fixing a hyperbolic metric on S, and after replacing all curves by their geodesic representatives, the lamination λ is obtained from any Hausdorff limit of the γ_i by deleting isolated leaves. We say that the γ_i *superconverge* to λ. By Proposition 3.15 we may assume that λ does not lie in the zero set of any compression body.

Suppose that $\pi_{\mathcal{H}(S)} \circ \gamma$ has finite diameter. By Theorem 4.1 there is a compression body V, a constant k, and a sequence of meridians $\partial D_i \in \mathcal{D}(V)$ so that $d_S(\gamma_i, \partial D_i) \leq k$. Kobayashi's Lemma [Kob88, Proposition 2.2] implies the ∂D_i also superconverge to λ. It follows that any accumulation point of the ∂D_i, taken in $\mathcal{PML}(S)$, is supported on λ. So, picking any measure of full support for λ, realizes λ as an element of $Z(V)$, contradicting our initial choice of λ. $\square$

6 Compression bodies of restricted topological type

Consider the poset $\mathcal{H}(S)$ of compression bodies, ordered by inclusion, as described above. We can take the quotient by the action of the mapping class group of the upper boundary surface S. This gives the poset of topological types of compression bodies, which we shall denote $\mathcal{H}_\top(S)$. Let A be a non-empty subset of $\mathcal{H}_\top(S)$. We say that A is *downwardly closed* if A is closed under passing to subcompression bodies. Define the *restricted compression body graph* $\mathcal{H}(S, A)$ to be the subposet of $\mathcal{H}(S)$ where we require all vertices to have topological type lying in A.

Theorem 6.1. *Let A be a downwardly closed connected non-empty subset of topological types of compression body. Then the restricted compression body graph $\mathcal{H}(S, A)$ is an infinite diameter Gromov hyperbolic metric space.*

Our methods apply in this generality, but in fact Theorem 6.1 also follows quickly from Theorem 1.1.

Proof of Theorem 6.1. If $A \subset B$ are downwardly closed subsets of $\mathcal{H}_\top(S)$ then the inclusion $\mathcal{H}(S, A) \subset \mathcal{H}(S, B)$ is simplicial and is coarsely onto. As a special case, $\mathcal{H}(S, A) \subset \mathcal{H}(S)$ is simplicial and coarsely onto. Thus $\mathcal{H}(S, A)$ has infinite diameter. $\square$

7 Loxodromics

Let G be a group acting by isometries on a Gromov hyperbolic space (X, d_X) with basepoint x_0. We say that an element $h \in G$ acts *loxodromically* if the translation length

$$\tau(h) = \lim_{n \to \infty} \tfrac{1}{n} d_X(x_0, h^n x_0)$$

is positive. This implies that h has a unique pair of fixed points $\{\lambda_h^+, \lambda_h^-\}$ in the Gromov boundary ∂X. Masur and Minsky [MM99] showed that for the action of the mapping class group on the curve complex, an element is loxodromic if and only if it is pseudo-Anosov. We say two loxodromic isometries h_1 and h_2 are *independent* if their pairs of fixed points $\{\lambda_{h_1}^+, \lambda_{h_1}^-\}$ and $\{\lambda_{h_2}^+, \lambda_{h_2}^-\}$ in the Gromov boundary ∂X are disjoint.

We remark that there are many pseudo-Anosov elements for which no power extends over any compression body. For example, if a power of g extends over a compression body, then some power of g preserves a rational subspace of first homology. However, generic elements of $\mathrm{Sp}(2g, \mathbb{Z})$ do not do this, see for example [DT06, Riv08]. We now give an alternate geometric argument to show the existence of pseudo-Anosov elements for which no power extends over a compression body.

Lemma 7.1. *Let S be a closed surface of genus at least two. Then there is a mapping class group element $h \in Mod(S)$ which acts loxodromically on the compression body graph $\mathcal{H}(S)$.*

We prove the following, more general result. Let G act by isometries on X, and let Y be a quasi-convex subset of X. We say that G *acts loxodromically* on (X, Y), if G contains a loxodromic element g whose quasi-axis is contained in a bounded neighbourhood of Y.

Lemma 7.2. *Let G be a group acting by isometries on a Gromov hyperbolic space X, and let $\mathcal{Y}$ be a collection of uniformly quasi-convex subsets of X. Furthermore, suppose that G acts with unbounded orbits on $X_{\mathcal{Y}}$, and there is a quasi-convex set $Y \in \mathcal{Y}$, such that G acts loxodromically on (X, Y). Then G contains an element which acts loxodromically on $X_{\mathcal{Y}}$.*

We now show that Lemma 7.2 implies Lemma 7.1.

Proof of Lemma 7.1. Let $G = \mathrm{Mod}(S)$ be the mapping class group acting on $X = \mathcal{C}(S)$, the complex of curves, which is Gromov hyperbolic. Let $\mathcal{Y}$ consist of the collection of disc sets of compression bodies, which is a collection of uniformly quasi-convex subsets of X. The space $X_{\mathcal{Y}} = \mathcal{H}(S)$ has infinite diameter by Theorem 1.1. The mapping class group acts coarsely transitively on $\mathcal{H}(S)$, and so in particular acts with unbounded orbits.

Let V be a handlebody and pick a pair of discs D and D' whose boundaries α and α' fill S. We define f to be the product of a right Dehn twist on α, followed by a left Dehn twist on α'. The mapping class group element f is pseudo-Anosov by work of Thurston [Thu88]. The element f extends over the compression body V, and the images of α under powers of f is a quasi-axis for f, which is contained in $\mathcal{D}(V)$, so f acts loxodromically on $\mathcal{D}(V)$.

Lemma 7.2 then implies that G contains an element which acts loxodromically on $\mathcal{H}(S)$. $\square$

The proof we present relies on the fact that (K, c)-quasi-geodesics in the curve complex $\mathcal{C}(S)$ project to reparameterized (K', c')-quasi-geodesics in the compression body graph $\mathcal{H}(S)$, where K' and c' depend only on K and c.

We will use the following properties of coarse negative curvature. We omit the proofs of Propositions 7.5, 7.6 and 7.8 below, as they are elementary exercises in coarse geometry.

We say a path γ is a (K, c, L)-*local quasi-geodesic*, if every subpath of γ of length L is a (K, c)-quasi-geodesic. The following theorem gives a classical "local to global" property for negative curvature.

Theorem 7.3. [CDP90, Chapter 3, Théorème 1.4, page 25] *Given constants δ, K and c, there are constants K', c' and L_0, such that for any $L \geqslant L_0$, any (K, c, L)-local quasi-geodesic in a δ-hyperbolic space X is a (K', c')-quasi-geodesic.*

In fact, we will make use of the following special case of this result. A *piecewise geodesic* is a path γ which is a concatenation of geodesic segments $\gamma_i = [x_i, x_{i+1}]$. We say a piecewise geodesic γ has *R-bounded Gromov products* if $(x_{i-1} \cdot x_{i+1})_{x_i} \leqslant R$ for all i.

Proposition 7.4. *Given constants δ and R, there is are constants L, K and c such that if γ is a piecewise geodesic in a δ-hyperbolic space, with R-bounded Gromov products, and $|\gamma_i| \geqslant L$ for all i, then γ is a (K, c)-quasi-geodesic.*

Proof. Consider a subpath of γ which is concatenation of two geodesic segments $[x_{i-1}, x_i]$ and $[x_i, x_{i+1}]$. Then by the definition of the Gromov product, this subpath is a $(1, 2(x_{i-1} \cdot x_{i+1})_{x_i})$-quasi-geodesic: that is, a $(1, 2R)$-quasi-geodesic. As each subpath of γ of length at most L is contained in at most 2 geodesic subsegments, γ is therefore a $(1, 2R, L)$-local quasi-geodesic. By Theorem 7.3, given $\delta, 1$ and $2R$, there are constants L, K and c, such that any $(1, 2R, L)$-local quasi-geodesic is a (K, c)-quasi-geodesic, as required. $\square$

We now record the following property of quasiconvex sets inside a hyperbolic space. This follows from thin triangles and we omit the proof.

Proposition 7.5. *Given constants δ and Q, there is a constant R_0, such that any constant $R \geqslant R_0$ has the following properties: let Y and Y' be Q-quasiconvex sets in a δ-hyperbolic space X. If the distance between Y and Y' is at least R, then the nearest point projection of Y to Y' has diameter at most R. Furthermore, if Y'' is a Q-quasiconvex set distance at least R from Y, and distance at most R from Y', then the distance between the nearest point projections of Y' and Y'' to Y is at most R.*

Let $\mathcal{Y}$ be a collection of Q-quasi-convex sets in a δ-hyperbolic space X, such that $X_{\mathcal{Y}}$ has infinite diameter.

Let $\mathcal{Z} = \{Z_i\}_{i \in N}$ be an ordered subcollection of the $\mathcal{Y}$, where N is a set of consecutive integers in $\mathbb{Z}$. If $N = \mathbb{Z}$, then we say that $\mathcal{Z}$ is *bi-infinite*. We say that $\mathcal{Z}$ is *L-well-separated* if $d_X(Z_i, Z_{i+1}) \geqslant L$ for all i, and furthermore the distance between the nearest point projections of Z_{i-1} and Z_{i+1} to Z_i is also at least L. Then $\mathcal{Z}$ determines a collection of piecewise geodesics as follows. For each Z_i, let p_i be a point in the nearest point projection of Z_{i-1} to Z_i, and let q_i be a point in the nearest point projection of Z_{i+1} to Z_i. Let $\gamma_{\mathcal{Z}}$ be a path formed from the concatenation of the geodesic segments $[p_i, q_i]$ and $[q_i, p_{i+1}]$. We will call such a path $\gamma_{\mathcal{Z}}$ a *$\mathcal{Z}$-piecewise geodesic*.

Proposition 7.6. *Given constants δ and Q, there is a constant R, such that for any Q-quasi-convex set Z in a δ-hyperbolic space X, and for any points x in X and z in Z, with p a nearest point in Z to x, then the Gromov product $(x \cdot z)_p \leqslant R$, where R depends only on δ and Q.*

We now show that an L-well separated collection $\mathcal{Z}$ of ordered Q-quasi-convex sets gives rise to a natural family of quasi-geodesics.

Proposition 7.7. *Given constants δ and Q, there are constants K, c and L_0, such that for any $L \geqslant L_0$, and any collection $\mathcal{Z}$ of L-well-separated ordered Q-quasi-convex sets in a δ-hyperbolic space X, any $\mathcal{Z}$-piecewise geodesic is a (K, c)-quasi-geodesic.*

Proof. By Proposition 7.5, there is a constant R_1, which only depends on δ and Q, such that if Z and Z' are two Q-quasi-convex sets distance at least R_1 apart, then the nearest point projection of Z to Z' has diameter at most R_1.

By Proposition 7.6, for any $x \in X$ and $z \in Z$, with p a nearest point in Z to x, then $(x \cdot z)_p \leqslant R_2$. For two adjacent segments $[p_i, q_i], [q_i, p_{i+1}]$ in a $\mathcal{Z}$-piecewise geodesic, q_i need not be the closest point on Z_i to p_i, but by Proposition 7.5, it is distance at most R_1 from the nearest point q_i' on Z_i to p_i. By the definition of the Gromov product, if $d_X(q_i, q_i') \leqslant R_1$, then the difference between the Gromov products $(p_i \cdot q_{i+1})_{q_i}$ and $(p_i \cdot q_{i+1})_{q_i'}$ is at most R_1, and so any $\mathcal{Z}$-piecewise geodesic has $(R_1 + R_2)$-bounded Gromov products. So, by Proposition 7.4, there are constants L, K and c, such that if every segment of a $\mathcal{Z}$-piecewise geodesic has length at least L, it is a (K, c)-quasi-geodesic. $\square$

We say that an ordered set $\mathcal{Z} = \{Z_i\}_{i \in \mathbb{Z}}$ of uniformly quasi-convex subsets of X is *(L, M)-$\mathcal{Y}$-separated*, if $\mathcal{Z}$ is L-well-separated in X, and $d_{X_{\mathcal{Y}}}(\pi(Z_i), \pi(Z_{i+1})) \geqslant M$ for all i.

Proposition 7.8. *Given constants δ, K and c there is a constant R, such that for any reparameterized (K, c)-quasi-geodesic $\gamma \colon \mathbb{R} \to X$ in a δ-hyperbolic space X, for any three numbers $r \leqslant s \leqslant t$, the Gromov product $(\gamma(r) \cdot \gamma(t))_{\gamma(s)} \leqslant R$.*

Proposition 7.9. *Given constants δ and Q, there are constants K, c, L_0 and M_0, such that for any collection $\mathcal{Y}$ of Q-quasi-convex sets, for any $L \geqslant L_0$*

and $M \geqslant M_0$, and any bi-infinite collection $\mathcal{Z}$ of (L,M)-$\mathcal{Y}$-separated sets, then the image of any $\mathcal{Z}$-piecewise geodesic in $X_{\mathcal{Y}}$ is a bi-infinite reparameterized (K,c)-quasi-geodesic.

Proof. By Proposition 7.7, given constants δ_1 and Q, there are constants K_1, c_1 and L_1 such that for any collection $\mathcal{Z}$ of L-well-separated Q-quasi-convex sets, with $L \geqslant L_1$, any $\mathcal{Z}$-piecewise geodesic γ is a (K_1, c_1)-quasi-geodesic in X. By Theorem 2.1 there are constants δ_2, K_2 and c_2 such that the image of γ in $X_{\mathcal{Y}}$ is a reparameterized (K_2, c_2)-quasi-geodesic in the δ_2-hyperbolic space $X_{\mathcal{Y}}$.

By Proposition 7.8, given constants δ_2, K_2 and c_2, there is a constant R, such that for any $i \leqslant j \leqslant k$, there is a bound on their Gromov product in $X_{\mathcal{Y}}$: that is, $(\pi(p_i) \cdot \pi(p_k))^{X_{\mathcal{Y}}}_{\pi(p_j)} \leqslant R$. Therefore, by Proposition 7.4, given constants δ_2 and R, there are constants K_2, c_2 and M such that as long as $d_{X_{\mathcal{Y}}}(\pi(p_i), \pi(p_{i+1})) \geqslant M$, the piecewise geodesic in $X_{\mathcal{Y}}$ formed from geodesic segments $[\pi(p_i), \pi(p_{i+1})]$ is a (K_2, c_2)-quasi-geodesic, and in particular is bi-infinite. So for any collection $\mathcal{Z}$ of (L,M)-$\mathcal{Y}$-separated sets, the image of any $\mathcal{Z}$-piecewise geodesic in $X_{\mathcal{Y}}$ is a reparameterized bi-infinite (K_2, c_2)-quasi-geodesic. □

We may now complete the proof of Lemma 7.2.

Proof (of Lemma 7.2). Let δ be the constant of hyperbolicity for X, and let Q be a constant such that all sets $Y \in \mathcal{Y}$ are Q-quasiconvex. Let f be an element of G which acts loxodromically on $Y \in \mathcal{Y}$. The isometry f acts elliptically on $X_{\mathcal{Y}}$, coarsely fixing Y. In particular, there is a constant A, depending only on δ and Q, such that $d_{X_{\mathcal{Y}}}(Y, f^\ell Y) \leqslant A$ for all $\ell \in \mathbb{Z}$.

Given δ and Q, let R_0 be the constant from Proposition 7.5, and choose $R \geqslant R_0 + A$. In particular, for any two sets Y and Y' in $\mathcal{Y}$, distance at least R apart in X, the nearest point projection of Y' to Y has diameter at most R, and if Y'' is another set in $\mathcal{Y}$, distance at least R from Y and at most R from Y', then the nearest point projections of Y' and Y'' to Y are distance at most R apart.

Consider a pair of numbers L and M with $M \geqslant L \geqslant R$. The group G acts with unbounded orbits on $X_{\mathcal{Y}}$, so for any such number M, there is a group element $g \in G$ such that $d_{X_{\mathcal{Y}}}(Y, gY) \geqslant M + A$. This implies that for all ℓ, $d_X(Y, gY) \geqslant M + A$, and equivalently, for all ℓ, $d_X(Y, g^{-1}Y) \geqslant M + A$. As $d_{X_{\mathcal{Y}}}(Y, f^\ell gY) = d_{X_{\mathcal{Y}}}(f^{-\ell}Y, gY)$ this implies that for all ℓ, $d_{X_{\mathcal{Y}}}(Y, f^\ell gY) \geqslant M$, and equivalently, that for all ℓ, $d_{X_{\mathcal{Y}}}(Y, g^{-1}f^{-\ell}Y) \geqslant M$. As the map from X to $X_{\mathcal{Y}}$ is 1-Lipschitz, this implies that $d_X(Y, f^\ell gY) \geqslant M$, and equivalently $d_X(Y, g^{-1}f^{-\ell}Y) \geqslant M$.

As we have chosen $M \geqslant R$, the nearest point projection of $f^\ell gY$ to Y has diameter at most R. Similarly, the nearest point projection of $g^{-1}f^{-\ell}Y$ to Y has diameter at most R. Furthermore, for all ℓ and ℓ', $d_X(g^{-1}f^{-\ell}Y, g^{-1}f^{-\ell'}Y) \leqslant A$, and so for all ℓ and ℓ', the nearest point projections of $g^{-1}f^{-\ell}Y$ and $g^{-1}f^{-\ell'}Y$ to Y are distance at most R apart. Therefore, as f acts loxodromically on Y, for any number $L \geqslant R$ there is a sufficiently large number ℓ such that $\pi_Y(g^{-1}f^{-\ell}Y)$ and $\pi_Y(f^\ell gY)$ are distance at least L apart. This is illustrated schematically in Figure 5.

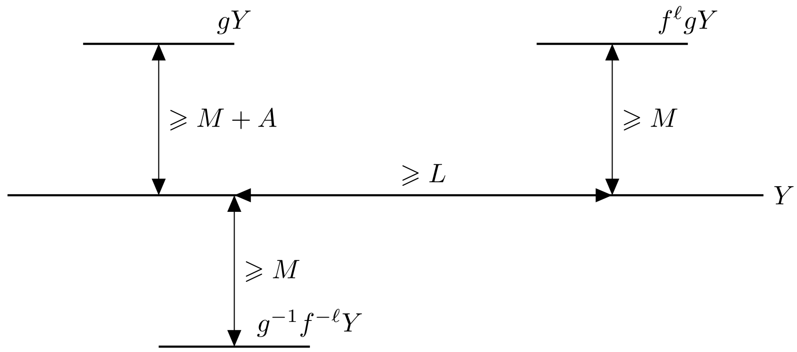


Figure 5: Nearest point projections of $g^{-1}Y$ and $f^\ell gY$ to Y in X.

Set $h = f^\ell g$. Then by Proposition 7.9, for L and M sufficiently large, the set $\mathcal{Z} = \{h^n Y\}_{n\in\mathbb{Z}}$ is a bi-infinite collection of (L, M)-$\mathcal{Y}$-separated sets, with the property that any $\mathcal{Z}$-piecewise geodesic quasi-axis for h projects to a bi-infinite quasi-axis for h in $X_\mathcal{Y}$, and so h acts loxodromically on $X_\mathcal{Y}$, as required. □

The following corollary implies that every subgroup in the Johnson filtration contains an element which acts loxodromically on $\mathcal{H}$.

Corollary 7.10. *Let X be a δ-hyperbolic space, and let $\mathcal{Y}$ be a collection of uniformly quasi-convex sets, such that the electrification $X_\mathcal{Y}$ has infinite diameter. Let G be a group which acts on X by isometries, and which contains an element which acts loxodromically on $X_\mathcal{Y}$. Let H be a normal subgroup of G, which contains an element which acts loxodromically on X. Then H contains an element which acts loxodromic on $X_\mathcal{Y}$.*

We will use the following property of independent loxodromics.

Proposition 7.11. *Let g and h be independent loxodromics on a δ-hyperbolic space X. Then there is a constant m, such that g^m and h^m freely generate a free group, all of whose non-trivial elements act loxodromically on X. Furthermore, the orbit map applied to this free group is a quasi-isometric embedding.*

Proof (of Corollary 7.10). Suppose that $g \in G$ acts loxodromically on $X_\mathcal{Y}$, and $h \in H$ acts loxodromically on X. If h acts loxodromically on $X_\mathcal{Y}$, then we are done. If not, then g and h are independent as loxodromics acting on X. We may choose $\{h^n x_0\}_{n\in\mathbb{Z}}$ as a quasi-axis γ_h for h. Note that the axis is quasi-convex in X.

Given constants L and M, we will show that there are positive integers l and m, such that if $f = g^l h^m g^{-l} h^m$, the ordered collection of sets $\mathcal{Z} = \{Z_n = f^n \gamma_h\}_{n\in\mathbb{N}}$ is a bi-infinite collection of uniformly quasi-convex (L, M)-$\mathcal{Y}$-separated sets in X. As $\mathcal{Z}$ consists of translates of a single quasi-convex set, the Z_i are uniformly quasi-convex.

For l sufficiently large, the nearest point projection of $f\gamma_h = g^l h^m g^{-l}\gamma_h$ to γ_h lies in a bounded neighbourhood of some vertex of γ_h, say x_0. Similarly, for l

and m sufficiently large, the nearest point projection of $f^{-1}\gamma_h$ lies in a bounded neighbourhood of $h^{-m}x_0$. By choosing m sufficiently large, this implies that

$$d_X(\pi_{\gamma_h}(f\gamma_h), \pi_{\gamma_h}(f^{-1}\gamma_h)) \geqslant L.$$

Then as $\mathcal{Z}$ is f-invariant, it is L-well-separated in X.

As g and h are independent, the nearest point projection of γ_h to γ_g has bounded diameter in X, so the nearest point projection of $\pi(\gamma_h)$ to $\pi(\gamma_g)$ also has bounded diameter in $X_{\mathcal{Y}}$. As g acts loxodromically on $X_{\mathcal{Y}}$, for any constant M there is a positive integer l such that $d_{X_{\mathcal{Y}}}(\gamma_h, f\gamma_h) \geqslant M$, and again as $\mathcal{Z}$ is f-invariant, this implies that $\mathcal{Z}$ is M-$\mathcal{Y}$-separated. Proposition 7.9 then implies that f acts loxodromically on $X_{\mathcal{Y}}$. □

We now show that if two pseudo-Anosov elements f and g act loxodromically on the compression body graph $\mathcal{H}(S)$, and act independently on $\mathcal{C}(S)$, then they also act independently on $\mathcal{H}(S)$.

Proposition 7.12. *Let X be a δ-hyperbolic space, and let $\mathcal{Y}$ be a collection of uniformly quasi-convex subsets of X. Let f and g act loxodromically on $X_{\mathcal{Y}}$, and let them be independent as loxodromics acting on X. Then they are independent loxodromics acting on $X_{\mathcal{Y}}$.*

Proof. Let $\lambda_f^{\pm}$ be the fixed points of f in ∂X, and let $\lambda_g^{\pm}$ be the fixed points of g in ∂X. As f and g act independently on X, all of these points are distinct. Let γ be a geodesic from λ_f^+ to λ_g^+. The image of γ under projection to $X_{\mathcal{Y}}$ is a reparameterized geodesic connecting the fixed points of f and g in $\partial X_{\mathcal{Y}}$ with at least one point in the interior of $X_{\mathcal{Y}}$. Thus $\pi \circ \gamma$ is a bi-infinite geodesic connecting the fixed points, so they are distinct. □

8 Applications

As an application of our methods, we give an alternative proof of a result of Biringer, Johnson and Minsky [BJM13, Theorem 1.1], and a slightly stronger version of a result of Lubotzky, Maher and Wu [LMW16, Theorem 2], using results of Maher and Tiozzo [MT17].

Theorem 8.1. *Let g be a pseudo-Anosov element of the mapping class group of a closed orientable surface. Then the stable lamination λ_g^+ of g is contained in the limit set for a compression body V if and only if the unstable lamination λ_g^- is contained in the limit set for V, and furthermore, there is a compression body $V' \subseteq V$ such that some power of g extends over V'.*

Proof. Suppose the stable lamination λ_g^+ for the pseudo-Anosov element g lies in the disc set of a compression body V_1. Let λ_g^- be the unstable lamination for g, and let γ be a geodesic axis for g in $\mathcal{C}(S)$. Choose a basepoint $\gamma(0)$ on γ, and assume that γ is parameterized by distance from γ_0 in $\mathcal{C}(S)$, with $\lim_{n\to\infty} \gamma(n) = \lambda_g^+$. Consider the sequence of compression bodies $W_i = g^{-i}V_1$.

Possibly after re-indexing the W_i, we may assume that the geodesic from $\gamma(0)$ to $\gamma(-i)$ is contained in k-neighbourhood of $\mathcal{D}(W_i)$, where k depends only on the constant of hyperbolicity and the quasi-convexity constants for the disc sets. We may therefore apply Theorem 4.2 above, which then implies that there is a compression body V_2, contained in infinitely many of the W_i. Furthermore, the geodesic ray from $\gamma(0)$ to $\lambda^-(g)$ is contained in a bounded neighbourhood of $\mathcal{D}(V_2)$. In particular, $V_2 \subseteq g^{k_1} V_1$ for some n_1, and we may assume that $k_1 \neq 0$ as V_2 is contained in infinitely many $g^i V_1$.

We may then repeat this procedure using positive powers of g, to produce a compression body V_3, contained in infinitely many $\mathcal{D}(g^i V_2)$, such that λ_g^+ is contained in the limit set of $\mathcal{D}(V_3)$. Iterating this procedure produces a descending chain of compression bodies $V_{n+1} \subseteq g^{k_n} V_n$, which has the property that the V_i have limit sets containing λ_g^+ if i is odd, and λ_g^- if i is even. As before, we may assume that the integers k_n are not zero. A descending chain of compression bodies must eventually stabilize with $V_n = g^{k_n} V_{n+1}$ for some n and $k_n \neq 0$. Therefore, there is a single compression body $V = V_n$, which contains both λ_g^+ and λ_g^-, as required. Finally, we observe that as $k_n \neq 0$, some power of g extends over this compression body. □

Lubotzky, Maher and Wu showed that a random walk on the mapping class group gives a hyperbolic manifold with a probability that tends to one exponentially quickly, assuming that the probability distribution μ generating the random walk is *complete*: that is, the limit set of the subgroup generated by the support of μ is dense in Thurston's boundary for the mapping class group, $\mathcal{PML}$. As the compression body graph is an infinite diameter Gromov hyperbolic space, we may apply the results of Maher and Tiozzo [MT17], to replace this hypothesis with the assumption that the support of μ contains a pair of independent loxodromic elements for the action of the subgroup on the compression body graph. We shall write w_n for a random walk of length n on the mapping class group of a closed orientable surface of genus g, generated by a probability distribution μ, and $M(w_n)$ for the corresponding *random Heegaard splitting*: that is the 3-manifold obtained by using the resulting mapping class group element w_n as the gluing map for a Heegaard splitting.

Proposition 8.2. *Every isometry of the compression body graph $\mathcal{H}$ is either elliptic or loxodromic.*

Proof. Biringer and Vlamis [BV17, Theorem 1.1] showed that the isometry group of $\mathcal{H}$ is equal to the mapping class group. If g is not pseudo-Anosov, then g acts elliptically on the curve complex, and hence acts elliptically on $\mathcal{H}$. Let g be a pseudo-Anosov element of the mapping class group. Let γ be an axis for g in the curve complex $\mathcal{C}(S)$, and let $\gamma(0)$ be a choice of basepoint for γ.

Suppose that g does not act loxodromically, it follows that there is a sequence of compression bodies V_i, whose nearest point projections to γ have diameters tending to infinity. Recall that the disc sets $\mathcal{D}(V_i)$ are uniformly quasi-convex. Thus there is a constant k, depending on the quasi-convexity constants of the disc sets and the hyperbolicity constant of the curve complex $\mathcal{C}(S)$, such that (after passing to a subsequence and re-indexing) the intersection of $\mathcal{D}(V_i)$ with a k-neighbourhood of γ has diameter at least i. By translating the disc sets by powers of g, and possibly passing to a further subsequence and re-indexing, we may assume that both $\gamma(0)$ and $\gamma(i)$ are distance at most k from $\mathcal{D}(g^{n_i} V_i)$. We may further relabel the disc sets and just write V_i for the given translate $g^{n_i} V_i$.

We may now apply our stability result, Theorem 4.2. This implies that there is a single compression body W contained in infinitely many of the V_i, and so in particular the positive limit point λ_g^+ of g is contained in the limit set of $\mathcal{D}(W)$. Thus there is a compression body W' such that some power of g extends over W'. Thus g acts elliptically, and we are done. $\square$

Theorem 8.3. *Let μ be a probability distribution on the mapping class group of a closed orientable surface of genus g, whose support has bounded image in the compression body graph $\mathcal{H}$, and which contains two independent pseudo-Anosov elements whose stable laminations do not lie in the limit set of any compression body. Then the probability that $M(w_n)$ is hyperbolic and of Heegaard genus g tends to one exponentially quickly.*

Proof. Maher and Tiozzo [MT17, Theorem 1.2] show that given a countable group G acting non-elementarily on a separable Gromov hyperbolic space (X, d_X), a finitely supported random walk on G has positive drift with exponential decay: that is, there are constants $L > 0, K \geqslant 0$, and $c < 1$ such that

$$\mathbb{P}\left[d_X(x_0, w_n x_0) \geqslant Ln\right] \geqslant 1 - Kc^n.$$

We may apply this result to the mapping class group G acting on the compression body graph $\mathcal{H}$. A subgroup of G acts *non-elementarily* on $\mathcal{H}$ if G contains two independent loxodromic elements. Geodesics in $\mathcal{C}(S)$ project to reparameterized quasi-geodesics in $\mathcal{H}$, so by Theorem 4.1, if the stable and unstable laminations of a pseudo-Anosov element h do not lie in the limit set of some compression body, then the image of the axis of h in $\mathcal{C}_{\mathcal{D}}(S)$ is a bi-infinite quasi-axis for the action of h on the compression body graph $\mathcal{H}$, and so in particular h acts loxodromically on $\mathcal{H}$. If h_1 and h_2 are two pseudo-Anosov elements, which act loxodromically on the compression body graph $\mathcal{H}$, and act independently on the curve complex $\mathcal{C}(S)$, then in fact they act independently on the compression body graph $\mathcal{H}$, as the geodesic from $\lambda_{h_1}^+$ to $\lambda_{h_2}^+$ in the curve complex $\mathcal{C}(S)$ projects to an reparameterized quasi-geodesic in $\mathcal{H}$, which has infinite diameter by Theorem 4.1. Finally, we observe that the compression body graph $\mathcal{H}$ is a countable simplicial complex, and so is separable: that is, has a countable dense subset, and so the hypotheses of [MT17, Theorem 1.2] are satisfied.

The *Hempel distance* of a Heegaard splitting $M(h)$ is the distance in the curve complex between the disc sets of the two handlebodies. This is bounded below by distance in the compression body graph $d_{\mathcal{H}}(x_0, w_n x_0)$. Since distance in the compression body graph has positive drift with exponential decay, so does Hempel distance. Finally, we observe that if the Hempel distance is at least three, then $M(w_n)$ is hyperbolic, by work of Hempel [Hem01, Corollaries 3.7 and 3.8], and Perelman's proof of Thurston's geometrization conjecture [MT07].

If the splitting distance is greater than $2g$, then the given Heegaard splitting is a minimal genus Heegaard splitting for $M(w_n)$, by work of Scharlemann and Tomova [ST06, page 594]. $\square$

A Train tracks

In this section we review the work of Masur and Minsky [MM04]. They prove the version of Proposition 3.13 we require, but we find it convenient to write down an argument which differs from theirs in certain details.

Let a and b be two essential simple closed curves in S in minimal position. A *subarc* of a simple closed curve is a closed connected subinterval. We will only every consider subarcs of a or b whose endpoints lie in $a \cap b$. Given a pair of curves a and b in minimal position, a *bicorn* is a simple closed curve, denoted by $c = (a_i, b_i)$, consisting of the union of one subarc a_i of a and one subarc b_i of b.

A subarc of a or b is *innermost* with respect to $a \cap b$ if its endpoints lie in $a \cap b$, and its interior is disjoint from $a \cap b$. We may abuse notation by referring to these innermost subarcs as components of either $a - b$ or $b - a$, though in fact we wish to include their endpoints. An innermost arc b_i of b with respect to $a \cap b$ is a *returning arc* if both endpoints lie on the same side of a in S. This is illustrated on the left hand side of Figure 6 below.

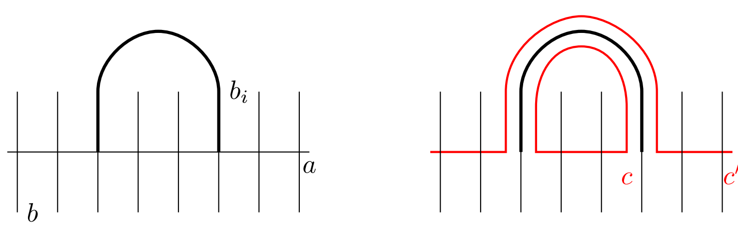


Figure 6: An innermost arc b_i of b forming a returning arc for a.

Given a simple closed curve a and a returning arc b_i, we may produce a new simple closed curve by *arc surgery* of a with respect to b_i. There are two possible bicorns, c and c', formed from the union of b_i with one of the two subarcs of a with endpoints ∂b_i. These are illustrated on the right hand side of Figure 6, where we have isotoped the replacement curves to be disjoint from b_i for clarity.

We say a bicorn is *returning* if b_i is a returning arc for a. We say a sequence of returning bicorns (a_i, b_i) is *nested* if $a_{i+1} \subset a_i$, and b_{i+1} is a returning arc for c_i, for all i. An adjacent pair in a sequence of returning bigons is illustrated in Figure 7.

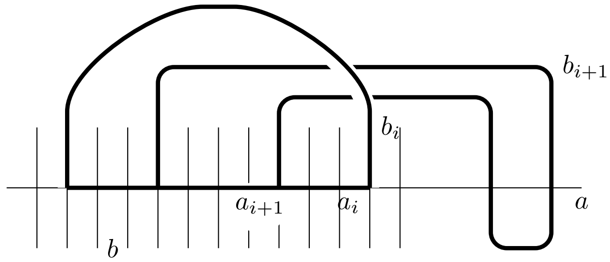


Figure 7: Nested bicorns.

For a nested bicorn sequence, each pair of adjacent bicorns c_i and c_{i+1} may be made disjoint after a small isotopy.

Let D and E be essential embedded discs in minimal position in a compression body. Then $a = \partial D$ and $b = \partial E$ is a pair of essential simple closed curves in minimal position. We say a disc F contained in $D \cup E$ is a *bicorn disc* if the boundary of the disc F is a bicorn in $a \cup b$.

We omit the proof of the following observation.

Proposition A.1. *Let D and E be essential embedded discs in minimal position in a compression body, and let F_i be a bicorn disc in $D \cup E$ with boundary $a_i \cup b_i$. Then there is an arc γ_i in $D \cap E$ such that F is the union of a subdisc D_i of D bounded by $a_i \cup \gamma_i$ and a subdisc E_i of E bounded by b_i and γ_i, which only intersect along γ_i.*

A *nested bicorn disc sequence* is a nested bicorn sequence in which every bicorn $c_i = (a_i, b_i)$ bounds a disc $F_i = (D_i, E_i)$, and furthermore $D_{i+1} \subset D_i$ for all i.

Two essential simple closed curves a and b in minimal position in S, and a subinterval $a_i \subset a$ with endpoints in $a \cap b$, determines a pre-train track τ_i' with a single switch, as follows: discard $a - a_i$, collapse a_i to a point, and smooth the tangent vectors as illustrated in Figure 8. The dashed line labelled a on the right hand side of Figure 8 is not part of the pre-train track τ_i'. Rather, it is drawn for comparison with the left hand diagram.

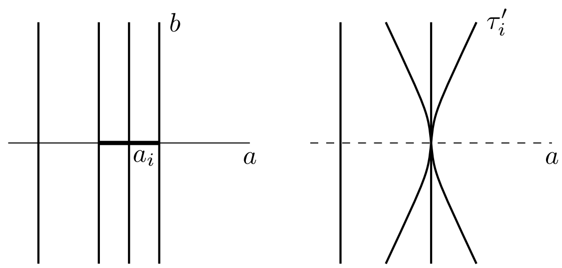


Figure 8: Smoothing intersections of a and b by collapsing a_i.

A pair of simple closed curves a and b in minimal position divide the surface S into a number of complementary regions, whose boundaries consist of alternating innermost subarcs of a and b. We say a complementary region is a *rectangle* if it is a disc, whose boundary consists of exactly four innermost subarcs.

Given a pre-train tack τ contained in a surface S, a *bigon collapse* is a homotopy of the surface, supported in a neighbourhood of a bigon, which maps the bigon to a single arc, as illustrated in Figure 9.

Figure 9: Collapsing a bigon.

We break the proof of Proposition 3.13 into the following three propositions.

We say a pre-track *collapses* to a train track if there is a sequence of bigon collapses which produces a train track. We say a bicorn (a_i, b_i) is *non-degenerate* if there is a non-rectangular component of $S-(a\cup b)$ whose boundary intersects $a - a_i$. Non-degeneracy allows us to collapse a pre-track to a track, as follows.

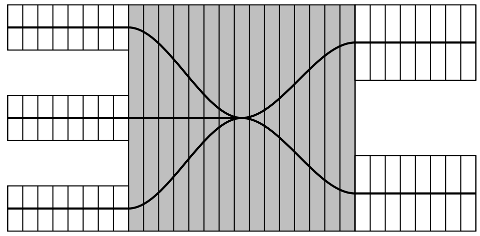
Proposition A.2. *Let a and b be simple closed curves in minimal position, and let (a_i, b_i) be a non-degenerate bicorn. Then the pre-track τ_i' determined by a_i and b bigon collapses to a train track τ_i. Furthermore, τ_i is switch dual to the bicorn c_i determined by (a_i, b_i).*

A bicorn sequence (a_i, b_i) is *non-degenerate* if every bicorn (a_i, b_i) is non-degenerate. If the initial bicorn (a_1, b_1) is non-degenerate, then this implies that every subsequent bicorn is non-degenerate.

Proposition A.3. *Let a and b be simple closed curves in minimal position, and let (a_i, b_i) be a collapsible nested bicorn sequence, and let τ_i be the corresponding train tracks. Then τ_i is a carrying sequence of train tracks: that is, $\tau_{i+1} \prec \tau_i$ for all i.*

Proposition A.4. *Let a and b be simple closed curves which bound discs in a compression body V. Then there is a collapsible nested bicorn disc sequence $\{F_i\}_{i=1}^n$ with bicorn boundaries $\{c_i\}_{i=1}^n$, as follows. If we define $c_0 = a$, then the sequence of simple closed curves $\{c_i\}_{i=0}^n$ is a disc surgery sequence connecting a and b.*

We now define a *rectangular tie neighbourhood* for a train track τ with a single switch. This is a regular neighbourhood of τ in the surface S which is foliated by intervals transverse to τ, with a decomposition as a union of foliated rectangles with disjoint interiors, with the following properties. There is a single rectangle containing the switch, which we shall call the *switch rectangle*. There is one rectangle for each branch, which we shall call a *branch rectangle*. The branch rectangles have disjoint closures.



The switch rectangle in the center is shaded.

Figure 10: A rectangular tie neighbourhood for a train track.

We define a *rectangular tie neighbourhood* for a pre-track with a single switch as above, except that we allow a single rectangle to contain multiple parallel branches. We observe that parallel branches in a single rectangle make up the boundaries of bigons in the pre-track, and these may be collapsed by a homotopy supported in the union of the switch rectangle and the rectangle containing the branches. In particular, if all bigon complementary regions are contained in the rectangular tie neighbourhood, then collapsing all bigons produces a train track, for which the rectangular tie neighbourhood of the pre-track is a rectangular tie neighbourhood for the train track.

We now define a collection of foliated rectangles determined by a and b. All of the rectangular tie neighbourhoods we construct will be subcollections of these rectangles. Isotope a and b into minimal position. For each intersection point x in $a \cap b$ we take a rectangular neighbourhood R_x, which we shall call a *vertex rectangle*. The rectangle has four *corners*, one in each of the quadrants formed by the local intersection of a and b, and alternating sides parallel to a and b. We foliate R_x by arcs parallel to $a \cap R_x$, so that the two sides of the rectangle parallel to a are leaves of the foliation. This is illustrated below in Figure 11.

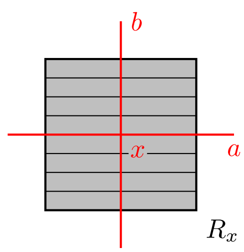


Figure 11: A rectangle R_x determined by a point $x \in a \cap b$.

Now suppose that α is a component of $a - b$, with endpoints x and y. We define a rectangle R_α, called an *a-rectangle*, as follows. It is a rectangle in S, containing $\alpha - (R_x \cup R_y)$, two of whose sides consist of the sides of R_x and R_y

which are parallel to b and intersect α. The other two sides consist of properly embedded arcs parallel to $\alpha - (R_x \cup R_y)$. We shall foliate this rectangle with arcs parallel to $a \cap R_\alpha$ such that the two sides parallel to α are leaves of the foliation. This is illustrated in Figure 12.

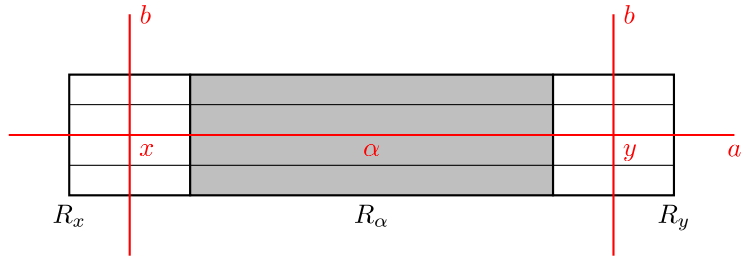


Figure 12: Rectangles determined by a component α of $a - b$.

We do the same for components β of $b - a$, but this time foliated by arcs crossing β exactly once, as illustrated in Figure 13. We shall call these rectangles *b-rectangles*.

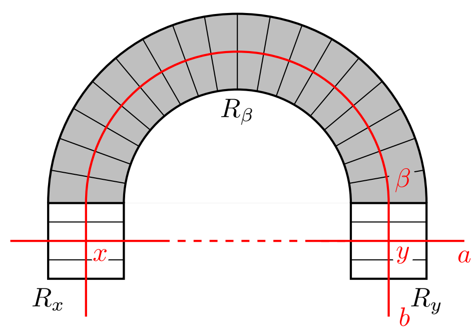


Figure 13: Rectangles determined by a component β of $b - a$.

Finally, suppose that f is a rectangle component of $S - (a \cup b)$. The *face rectangle* R_f is a foliated rectangle lying inside f, whose sides consist of the four sides of the a- and b-rectangles which meet f. The foliation consists of arcs parallel to a, such that the two a-sides of the rectangle are leaves of the foliation. This is illustrated below in Figure 14.

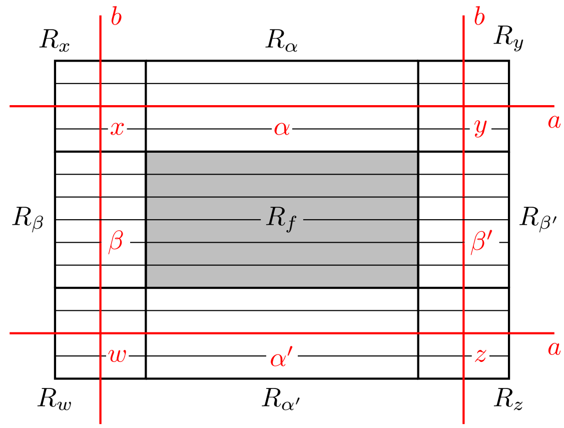


Figure 14: A face rectangle determined by a rectangular face f of $S - (a \cup b)$.

We shall denote the resulting foliation of this subset of S by $\mathcal{F}$. This foliates all of S except for the (slightly shrunken) non-rectangular regions of $S-(a\cup b)$.

We now prove Proposition A.2. We will construct a foliated region in S which is a union of rectangles, and show that it is a tie neighbourhood for a train track with a single switch.

Proof (of Proposition A.2). The foliated region $\mathcal{F}$ is a union of foliated rectangles. We will build a rectangular tie neighbourhood $\mathcal{F}_i$ for the pre-track τ_i' which will consist of a subcollection of these rectangles, and which may contain branch rectangles with multiple edges. There will be no complementary regions of $\mathcal{F}_i$ which are rectangles, so the pre-track τ_i' will collapse to a train track τ_i.

The foliated region $\mathcal{F}_i$ consists of all vertex rectangles and all b-rectangles, together with all the face rectangles which are contained in rectangular components of $S-(a_i \cup b)$, as well as all a-rectangles adjacent to an included face rectangle.

Let a_i^+ be the maximal subarc of a with endpoints in $a \cap b$ contained in the connected component of $\mathcal{F}_i \cap a$ containing a_i. In particular $a_i \subseteq a_i^+$. As the bicorn (a_i, b_i) is non-degenerate, there is at least one non-rectangle region of $S-(a\cup b)$ with a boundary edge in $a - a_i$. Any non-rectangle region of $S-(a\cup b)$ is contained in a non-rectangle region of $S-(a_i\cup b)$, so a_i^+ does not consist of all of a_i, and therefore is an interval.

The union of the vertex rectangles and the a-rectangles in $\mathcal{F}_i$ meeting a_i^+ is a regular neighbourhood of a_i^+, and is a foliated rectangle containing the unique switch of the pre-track τ_i. We shall denote this rectangle by R_i^+. This will be the switch rectangle for the rectangular tie neighbourhood.

We must now show that all components of $\mathcal{F}_i - R_i^+$ are foliated rectangles (perhaps containing multiple branches of τ_i) and that no components of $S-\mathcal{F}_i$ is a rectangle.

If a component of $\mathcal{F}_i - R_i^+$ has no face rectangles, then it is a union of b-edge rectangles and vertex rectangles. Let $\mathcal{B}$ be the union of the vertex rectangles and the b-rectangles; so $\mathcal{B}$ is a regular neighbourhood of b. The intersection $\mathcal{B} \cap R_i^+$ consists of a subset of the vertex rectangles. Thus $\mathcal{B} - R_i^+$ is a union of rectangles with disjoint closures. We will refer to the components of $\mathcal{B} - R_i^+$ as *b-strips*. Each b-strip is a rectangle whose boundary consists of four edges: two parallel to a, and contained in ∂R_i^+, and the other two parallel to b, and parallel to properly embedded arcs in the complement of R_i^+. We say two b-strips are *parallel* if their b-parallel edges cobound a rectangle in $S - R_i^+$.

We say two face rectangles are *vertically adjacent* if they border a common a-rectangle not in R_i^+. We say a maximal collection of vertically adjacent rectangles, together with the a-rectangles between them, is a *face strip*. A face strip is a rectangle whose boundary consists of four edges, two parallel to a and contained in ∂R_i^+, and two parallel to b, and are properly embedded parallel arcs in the complement of R_i^+. We say two face strips are *parallel* if the b-parallel edges are parallel arcs in the complement of R_i^+.

We say a face strip is *parallel* to a b-strip if their b-parallel edges cobound a rectangle in $S - R_i^+$. We observe that a face strip is adjacent to two b-strips, one on each side, and all three strips are parallel to each other.

A component of $\mathcal{F}_i - R_i^+$ which contains a face rectangle is a union of face strips and b-strips. Each pair of face strips is separated by a b-strip, and the observation in the paragraph above ensures that all of the strips are parallel.

We now verify that all of the components of $\mathcal{F}_i - R_i^+$ are disjoint. Any two components consisting only of b-strips are disjoint. Let R_1 and R_2 be a pair of components of $\mathcal{F}_i - R_i^+$ which contain a corner in common. Each corner lies in the boundary of a face rectangle, a b-rectangle, an a-rectangle and a vertex rectangle. Suppose the face rectangle lies in one of the components, say R_1. Then the b-rectangle also lies in R_1. If the vertex and a-rectangles do not lie in R_1, then they lie in R_i^+, so there can be no rectangle R_2 intersecting R_1, a contradiction. So neither R_1 nor R_2 contain the face rectangle. If the a-rectangle lies in R_1, then so must the face rectangle beside it, so the a-rectangle is also not contained in either R_1 or R_2. If R_1 contains the b-rectangle, then it either contains the vertex rectangle, or the vertex rectangle lies in a_i^+. In either case, none of the other rectangles can lie in R_2, a contradiction. □

Proposition A.3 follows from work of Penner and Harer [PH92].

Proof (of Proposition A.3). Any closed train route in τ_{i+1} is a union of arcs of b starting and ending at a_{i+1}. As $a_{i+1} \subset a_i$, it is also a train route in τ_i. The train track τ_{i+1} may then be obtained from τ_i by splitting and shifting, by Penner and Harer [PH92, Theorem 2.4.1]. □

Before we prove Proposition A.4 we need the following observation.

Proposition A.5. *Suppose that a and b are essential simple closed curves in minimal position. Then any bicorn contained in $a \cup b$ is an essential simple closed curve.*

Proposition A.5 is standard and we omit the proof.

Proof (of Proposition A.4). Let $a = \partial D$ and $b = \partial E$. We first show how to choose an initial non-degenerate bicorn (a_1, b_1). By an Euler characteristic argument, there must be at least one complementary region of $a \cup b$ which is not a rectangle.

Let α be a component of $a - b$ which lies in the boundary of one of the non-rectangular regions of the complement of $a \cup b$. Let $b_1 \subset b$ be a choice of returning arc determined by an outermost disc E_1 in E, and let a_1 be the subinterval of a:

- with the same endpoints as b_1,
- which is disjoint from α.

The bicorn c_1 determined by (a_1, b_1) bounds a disc corresponding to disc surgery of D along E_1, and we shall denote this disc by F_1. This disc F_1 is essential by Proposition A.5.

Now suppose we have constructed the i-th nested bicorn disc $F_i = (D_i, E_i)$ with bicorn boundary $c_i = (a_i, b_i)$. The intersection $E \cap F_i$ is equal to $E_i \cup (E \cap D_i)$. Choose E_{i+1} to be an outermost disc of E with respect to $E \cap D_i$, which is not E_i. Then $\gamma_{i+1} = E_{i+1} \cap D_i$ bounds a disc in D_i which we shall choose to be D_{i+1}. We shall set $F_{i+1} = D_{i+1} \cup E_{i+1}$, which has bicorn boundary $c_{i+1} = (a_{i+1}, b_{i+1})$, where $a_{i+1} \subset a_i$ to be $D_{i+1} \cap a$ and b_{i+1} to be $E_{i+1} \cap b$. The disc F_{i+1} is essential by Proposition A.5, and furthermore, is obtained from disc surgery of F_i along E_{i+1}. We have therefore constructed the next disc in the nested bicorn disc sequence, as required. □

References

[Ack15] Robert Ackermann, *An alternative approach to extending pseudo-Anosovs over compression bodies*, Algebr. Geom. Topol. **15** (2015), no. 4, 2383–2391.

[AF19] Ian Agol and Michael Freedman, *Embedding heegaard decompositions* (2019), available at `arXiv:1906.03244`.

[BJM13] Ian Biringer, Jesse Johnson, and Yair Minsky, *Extending pseudo-Anosov maps into compression bodies*, J. Topol. **6** (2013), no. 4, 1019–1042.

[BV17] Ian Biringer and Nicholas G. Vlamis, *Automorphisms of the compression body graph*, J. Lond. Math. Soc. (2) **95** (2017), no. 1, 94–114.

[Bow12] Brian H. Bowditch, *Relatively hyperbolic groups*, Internat. J. Algebra Comput. **22** (2012), no. 3, 1250016, 66.

[BP14] Stephan D. Burton and Jessica S. Purcell, *Geodesic systems of tunnels in hyperbolic 3-manifolds*, Algebr. Geom. Topol. **14** (2014), no. 2, 925–952.

[CL85] A. J. Casson and D. D. Long, *Algorithmic compression of surface automorphisms*, Invent. Math. **81** (1985), no. 2, 295–303.

[CDP90] M. Coornaert, T. Delzant, and A. Papadopoulos, *Géométrie et théorie des groupes*, Lecture Notes in Mathematics, vol. 1441, Springer-Verlag, Berlin, 1990 (French).

[DP17] Vinh Dang and Jessica Purcell, *Cusp shape and tunnel number* (2017), available at `arXiv:1711.03693`.

[DT17] Spencer Dowdall and Samuel J. Taylor, *The co-surface graph and the geometry of hyperbolic free group extensions*, J. Topol. **10** (2017), no. 2, 447–482.

[DT06] Nathan M. Dunfield and William P. Thurston, *Finite covers of random 3-manifolds*, Invent. Math. **166** (2006), no. 3, 457–521.

[LM12] Martin Lustig and Yoav Moriah, *Are large distance Heegaard splittings generic?*, J. Reine Angew. Math. **670** (2012), 93–119. With an appendix by Vaibhav Gadre.

[Ham06] Ursula Hamenstädt, *Train tracks and the Gromov boundary of the complex of curves*, Spaces of Kleinian groups, London Math. Soc. Lecture Note Ser., vol. 329, Cambridge Univ. Press, Cambridge, 2006, pp. 187–207.

[Hem01] John Hempel, *3-manifolds as viewed from the curve complex*, Topology **40** (2001), no. 3, 631–657.

[KM96] Vadim A. Kaimanovich and Howard A. Masur, *The Poisson boundary of the mapping class group*, Invent. Math. **125** (1996), no. 2, 221–264.

[KR14] Ilya Kapovich and Kasra Rafi, *On hyperbolicity of free splitting and free factor complexes*, Groups Geom. Dyn. **8** (2014), no. 2, 391–414.

[Kla] E. Klarreich, *The boundary at infinity of the curve complex and the relative Teichmüller space.*

[Kob88] Tsuyoshi Kobayashi, *Heights of simple loops and pseudo-Anosov homeomorphisms*, Braids (Santa Cruz, CA, 1986), Contemp. Math., vol. 78, Amer. Math. Soc., Providence, RI, 1988, pp. 327–338.

[LMW16] Alexander Lubotzky, Joseph Maher, and Conan Wu, *Random Methods in 3-Manifold Theory*, Tr. Mat. Inst. Steklova **292** (2016), no. Algebra, Geometriya i Teoriya Chisel, 124–148.

[MW17] Jiming Ma and Jiajun Wang, *Quasi-homomorphisms on mapping class groups vanishing on a handlebody group* (2017).

[Mah10] Joseph Maher, *Random Heegaard splittings*, J. Topol. **3** (2010), no. 4, 997–1025.

[MT17] Joseph Maher and Giulio Tiozzo, *Random walks on weakly hyperbolic groups*, Journal für die reine und angewandte Mathematik (Crelle's Journal), to appear (2017).

[Mas86] Howard Masur, *Measured foliations and handlebodies*, Ergodic Theory Dynam. Systems **6** (1986), no. 1, 99–116.

[MM99] Howard A. Masur and Yair N. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math. **138** (1999), no. 1, 103–149.

[MM00] H. A. Masur and Y. N. Minsky, *Geometry of the complex of curves. II. Hierarchical structure*, Geom. Funct. Anal. **10** (2000), no. 4, 902–974.

[MM04] Howard A. Masur and Yair N. Minsky, *Quasiconvexity in the curve complex*, In the tradition of Ahlfors and Bers, III, Contemp. Math., vol. 355, Amer. Math. Soc., Providence, RI, 2004, pp. 309–320.

[MMS12] Howard Masur, Lee Mosher, and Saul Schleimer, *On train-track splitting sequences*, Duke Math. J. **161** (2012), no. 9, 1613–1656.

[MS13] Howard Masur and Saul Schleimer, *The geometry of the disk complex*, J. Amer. Math. Soc. **26** (2013), no. 1, 1–62.

[MT07] John Morgan and Gang Tian, *Ricci flow and the Poincaré conjecture*, Clay Mathematics Monographs, vol. 3, American Mathematical Society, Providence, RI, 2007.

[PH92] R. C. Penner and J. L. Harer, *Combinatorics of train tracks*, Annals of Mathematics Studies, vol. 125, Princeton University Press, Princeton, NJ, 1992.

[Riv08] Igor Rivin, *Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms*, Duke Math. J. **142** (2008), no. 2, 353–379.

[ST06] Martin Scharlemann and Maggy Tomova, *Alternate Heegaard genus bounds distance*, Geom. Topol. **10** (2006), 593–617.

[Thu88] William P. Thurston, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bull. Amer. Math. Soc. (N.S.) **19** (1988), no. 2, 417–431.

Joseph Maher
CUNY College of Staten Island and CUNY Graduate Center
joseph.maher@csi.cuny.edu

Saul Schleimer
University of Warwick
s.schleimer@warwick.ac.uk