A Recipe for Short-Word Pseudo-Anosovs, and More

by

Johanna Ceres Isabel Mangahas Kutluhan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mathematics)
in The University of Michigan
2010

Doctoral Committee:

    Professor Richard D. Canary, Co-Chair
    Assistant Professor Juan Souto, Co-Chair
    Professor Mario Bonk
    Professor Ralf J. Spatzier
    Assistant Professor Vilma M. Mesa

for Grandmo

ACKNOWLEDGMENTS

Dick Canary and Juan Souto, my thesis advisors, deserve enormous gratitude for tolerating their student, as well as for their insight, guidance, and support. I also have great appreciation for the intellect and generosity of the many mathematicians in my field. In particular, conversations with Chris Leininger brought me so much of the knowledge necessary for this work. Yair Minsky convinced me that the hands-on proof of Behrstock's Lemma would work for the annulus. Lee Mosher explained such a nice, unified view of mapping class group results, around the notion of an active subsurface. I would be nowhere without Nikolai Ivanov's extensive investigation of mapping class subgroups. The work of Koji Fujiwara gives my own contributions relevance.

Mathematical thank-you's spill easily into personal gratitude, and then I have too many to credit. So let me save the last mention for my husband Çağatay, whose love and encouragement never waver.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

Introduction

For an oriented topological surface S, a *mapping class* is the isotopy class of an orientation-preserving homeomorphism from S to itself. These constitute a discrete, infinite group known as $\mathrm{Mod}(S)$, the *mapping class group* of S. The group $\mathrm{Mod}(S)$ acts on the set of isotopy classes of homotopically nontrivial simple closed curves on S. If all elements of a subgroup fix no common family of curves, then the group contains a *pseudo-Anosov*, that is, a single element which itself fixes no finite family of curves [19]. This dissertation answers in the affirmative Fujiwara's question of whether one can always find a "short-word" pseudo-Anosov (Question 3.4 in [9]). Where Σ generates the group G, let Σ-length denote the length of an element of G in the word metric induced by Σ. The affirmative statement is:

Theorem I.1. *There exists a constant $K = K(S)$ with the following property. Suppose $G < \mathrm{Mod}(S)$ is finitely generated by Σ and contains a pseudo-Anosov. Then G contains a pseudo-Anosov with Σ-length less than K.*

The proof provides an explicit construction of pseudo-Anosovs from arbitrary non-pseudo-Anosov elements. In fact, it addresses a broader question. Roughly speaking, a pseudo-Anosov requires the whole surface for its support. The other elements are called reducible, because they allow one to "reduce" the surface in question.

That is, reducible mapping classes have powers which fix proper subsurfaces and act trivially or as a pseudo-Anosov on those subsurfaces. Mosher has described a unifying approach to both types [31], associating to a mapping class f what he calls its active subsurface $\mathcal{A}(f)$. For the sake of introduction one may think of $\mathcal{A}(f)$ as the support of f (really, it is the smallest subsurface supporting some power of f), remember that these subsurfaces and thus their inclusion are defined up to isotopy, and observe that f is pseudo-Anosov exactly when $\mathcal{A}(f) = S$. Then several foundational mapping class group theorems—including the Tits alternative for $\mathrm{Mod}(S)$ [19, 29], and subgroup structure results from [4] and [18]—elegantly derive from what Mosher coined the Omnibus Subgroup Theorem[1]: given a group $G < \mathrm{Mod}(S)$ there exists $f \in G$ such that for all $g \in G$, $\mathcal{A}(g) \subset \mathcal{A}(f)$. Call such an f *full-support for G*. In this paper we actually prove the following theorem, which quantifies the Omnibus Theorem and includes Theorem I.1 as a special case.

Theorem III.1. (Main Theorem.) *There exists a constant $K = K(S)$ such that, for any finite subset $\Sigma \subset \mathrm{Mod}(S)$, one may find f full-support for $\langle \Sigma \rangle$ with Σ-length less than K.*

The proof of Theorem III.1 spells out short pseudo-Anosovs explicitly, with the following core construction, concerning special pairs of pure reducible mapping classes we will call *sufficiently different*. These are pure mapping classes a and b with pseudo-Anosov restrictions to proper subsurfaces A and B respectively, such that A and B together fill S, meaning that each curve on S has essential intersection with either A or B. The proposition also identifies subgroups that are convex cocompact, a property defined for mapping class subgroups by Farb and Mosher, in analogy to

---

[1]We should clarify that Mosher formulated the Omnibus Theorem to yield as corollaries results that we need to prove Theorem III.1. Mosher wanted streamlined proofs for transposing to the key of $\mathrm{Out}(F_n)$ (see [14] for progress). Content to use the old results as-is, we define active subsurfaces slightly differently, take the Omnibus Theorem as the corollary, and reap the benefits of its perspective and terminology.

convex cocompact Kleinian groups [8]. Convex cocompactness is proven for interest, and is not necessary for the main theorem.

Proposition I.2. *There exists a constant $Q = Q(S)$ with the following property. Suppose a and b are sufficiently different pure reducible mapping classes. Then for any $n, m \geq Q$, every nontrivial element of $G = \langle a^n, b^m \rangle$ is pseudo-Anosov except those conjugate to powers of a^n or b^m. Furthermore, G is a rank two free group, and all of its finitely generated all-pseudo-Anosov subgroups are convex cocompact.*

In [26], the author considers a more general condition for pairs of pure reducible mapping classes, and finds Q such that G as above is a rank two free group, but need not contain pseudo-Anosovs (and therefore need not be convex cocompact). A more relevant comparison is Thurston's theorem providing the first concrete examples of pseudo-Anosovs [36] (see also [33]). He proved that if a and b are Dehn twists about filling curves, then one can find an affine structure on S inducing an embedding of $\langle a, b \rangle$ into $PSL(2, \mathbb{R})$ under which hyperbolic elements of $PSL(2, \mathbb{R})$ correspond to pseudo-Anosovs in $\langle a, b \rangle$. In particular, every nontrivial element of the free semi-group generated by a and b^{-1} is pseudo-Anosov. More recently, Hamidi-Tehrani [13] classified all subgroups generated by a pair of positive Dehn multi-twists. In partic-ular, if α and β are multicurves whose union fills S, and a and b are compositions of positive powers of Dehn twists about components of α and β respectively, then except for finitely many pairs n, m, $\langle a^n, b^m \rangle$ is a rank-two free group whose only reducible elements are those conjugate to powers of a or b (see also [17]). In a dif-ferent light, one can consider Proposition I.2 a companion to a theorem of Fujiwara that generates convex cocompact free groups using bounded powers of independent pseudo-Anosovs; this theorem appears in Section 3.3.2 as Theorem III.2.

A separate motivation for Theorem I.1 is its utility for quantifying the mapping

class group version of the Tits alternative: a mapping class subgroup that is not virtually abelian contains a rank-two free group. This was proved independently by Ivanov and McCarthy [19], [29]. Theorem I.1 plus the aforementioned theorem of Fujiwara lead to the following "strong" Tits alternative for $\mathrm{Mod}(S)$:

Theorem IV.1. *Suppose $G < \mathrm{Mod}(S)$ is finitely generated and not virtually abelian. There exists a constant $w = w(S)$ such that, for any finite set Σ generating G, some pair of elements with Σ-length less than w generates a rank two free group.*

The constant w necessarily depends on S; we include this result as Theorem IV.4, proven jointly by E. Breuillard and the author.

The original motivation for our main theorems is a corollary of Theorem IV.1. It says that, for subgroups of $\mathrm{Mod}(S)$, exponential growth implies uniform exponential growth, and furthermore their minimal growth rates are uniformly bounded:

Corollary IV.2. *There exists a positive lower bound, depending only on S, for the minimal exponential growth rate of any subgroup of $\mathrm{Mod}(S)$ with exponential growth.*

Also corollary to Theorem IV.1 is the fact that, for a simple random walk on a non-virtually-abelian mapping class subgroup, return probability decays exponentially with a rate depending only on S and the initial support of the walk. We include this as Corollary IV.3.

Chapter II presents the background for Theorems I.1 & III.1 and Proposition I.2; Chapter III provides the proofs. Chapter IV elaborates on Theorems IV.1 & IV.4 and Corollaries IV.2 & IV.3; except for the derivation of Theorem IV.1, the material of Chapter IV appeared previously in [26].

CHAPTER II

Background

This section consists of five parts. The first presents basic definitions around the mapping class group and curve complex; the second relates the curve complexes of a surface and its subsurfaces. Section 2.3 presents a different view of subsurfaces, which facilitates Section 2.4 on the case for understanding reducible mapping classes as subsurface pseudo-Anosovs. Finally, Section 2.5 recalls powerful curve complex tools with which we rephrase the classification of elements of $\text{Mod}(S)$.

2.1 Mapping class group and curve complex

Throughout, we consider only oriented surfaces whose genus g and number of boundary components p are finite. Define the *complexity* $\xi(S)$ of a surface S by $\xi = 3g + p - 3$. We neglect the case where ξ is -2 or zero, which means S is a disk, a closed torus, or a pair of pants, because these are never subsurfaces of interest, as explained in Section 2.2. Annuli, for which $\xi = -1$, feature throughout, but primarily as subsurfaces of higher-complexity surfaces. Let us first assume $\xi \geq 1$, and address the annulus case after. Note that our definitions will not distinguish between boundary components and punctures, except on an annulus. One may find in [8] a discussion on variant definitions of the mapping class group.

Given a surface S, its *mapping class group* $\text{Mod}(S)$ is the discrete group of

orientation-preserving homeomorphisms from S to itself that setwise fix components of ∂S. Much (arguably, everything) about $\mathrm{Mod}(S)$ appears in its action on the isotopy classes of those simple closed curves on S that are *essential*: neither homotopically trivial nor boundary-parallel. Let us call these classes *curves* for short. *Pseudo-Anosov* mapping classes are those that fix no finite family of curves; necessarily these have infinite order. Among the rest, we distinguish those that have finite order, and call the remaining *reducible*.

The *intersection* of two curves α and β, written $i(\alpha, \beta)$, is the minimal number of points in $\alpha' \cap \beta'$, where α' ranges over all representatives of the isotopy class denoted by α and likewise β' ranges over representatives of β. Often, we will use the same notation for a curve as an isotopy class and for a representative path on the surface. For specificity and guaranteed minimal intersections, one may represent all curves by closed geodesics with respect to any pre-ordained hyperbolic metric on S. Two curves *fill* a surface if every curve of that surface intersects at least one of them. *Disjoint* curves have zero intersection. All of these definitions extend to *multicurves*, our term for sets of pairwise disjoint curves.

Now let us upgrade the set action of $\mathrm{Mod}(S)$ on curves to a simplicial action on the *curve complex* of S, denoted $\mathcal{C}(S)$. Because $\mathcal{C}(S)$ is a *flag complex*, its data reside entirely in the one-skeleton $\mathcal{C}_1(S)$: higher-dimensional simplices appear whenever the low-dimensional simplices allow it. Thus it suffices to consider only the graph $\mathcal{C}_1(S)$, although we usually refer to the full complex out of habit.

Curves on S comprise the vertex set $\mathcal{C}_0(S)$, and an intersection rule determines the edges of $\mathcal{C}_1(S)$. For a surface with complexity $\xi > 1$, edges join vertices representing disjoint curves. Therefore n-simplices correspond to multicurves with n distinct components, and ξ gives the dimension of $\mathcal{C}(S)$. When $\xi = 1$, S is a punctured torus or

four-punctured sphere. Because on these any two distinct curves intersect, we modify the previous definition so that edges join vertices representing curves which intersect minimally for distinct curves on that surface—that is, once for the punctured torus and twice for the four-punctured sphere. In both cases $\mathcal{C}(S)$ corresponds to the Farey tesellation of the upper half plane (vertices sit on rationals corresponding to slopes of curves on the torus).

Give $\mathcal{C}_1(S)$ the path metric where edges have unit length; this extends simplicially to the full complex. For any two curves α and β, let $\mathrm{d}(\alpha, \beta)$ denote their distance in $\mathcal{C}_1(S)$. Immediately one may observe that $\mathcal{C}(S)$ is locally infinite. It is not obvious that $\mathcal{C}(S)$ is connected, but the proof is elementary [15]. Furthermore, it has infinite diameter [27]. A deep theorem of Masur and Minsky states that $\mathcal{C}(S)$ is δ-hyperbolic, meaning that for some δ, every edge of any geodesic triangle lives in the δ-neighborhood of the other two edges [27].

Now suppose S is the annulus $S^1 \times [0, 1]$. We modify the definition of the mapping class group to require that homeomorphisms and isotopy fix ∂S pointwise. Parameterizing S^1 by angle θ and S by (θ, t), the *Dehn twist* on S maps (θ, t) to $(\theta + 2\pi t, t)$. The cyclic group generated by this twist is the entire mapping class group of the annulus. In this group, let us consider all non-trivial elements pseudo-Anosov, for reasons clarified in Section 2.5. To define $\mathcal{C}(S)$, we contend with the fact that an annulus contains no essential curves. Instead, each vertex of $\mathcal{C}_0(S)$ corresponds to the isotopy class of an arc connecting the two boundary components, again requiring isotopy fix boundary pointwise. Edges connect vertices that represent arcs with disjoint interiors. Although this complex is locally uncountable, it is not difficult to understand: the distance between two distinct vertices equals one plus the minimal number of interior points at which their representative arcs intersect. Again, $\mathcal{C}(S)$ is

connected, infinite diameter, and δ-hyperbolic—in fact, it is quasi-isometric to the real line, and $\mathrm{Mod}(S)$ acts on it by translation. Note that our "curve complex" for the annulus is more accurately called an *arc* complex. Generally, we aim to minimize the distinction between annular subsurfaces and subsurfaces with $\xi \geq 1$; for extended treatment of curve complexes and arc complexes, see [28].

2.2 Subsurface projection

Let us relate the curve complex of S to that of S', where $\xi(S) \geq 1$ and S' is an "interesting" subsurface of S. Here, a subsurface is defined only up to isotopy, and assumed *essential*, meaning its boundary curves are either essential in S or shared with ∂S. This rules out the disk. We also disregard pants, which have trivial mapping class group. For the remainder of Section 2.2, we assume S' is connected; furthermore, either $\xi(S') \geq 1$ or S' is an annular neighborhood of a curve in S.

For ease of exposition and the convenience of considering $\partial S'$ a multicurve in S, let us make a convention that $\partial S'$ refers only to those boundary curves essential in S. In later sections we consider multiple, possibly nested subsurfaces, but these are always implicitly or explicitly contained in some largest surface S.

Except when S' is an annulus, it is clear one can embed $\mathcal{C}(S')$, or at least its vertex set, in $\mathcal{C}(S)$, but we seek a map in the opposite direction. One can associate curves on the surface to curves on a subsurface via subsurface projection, a notion appearing in [18, 19], expanded in [28], and recapitulated here. In what follows, we define the projection map $\pi_{S'}$ from $\mathcal{C}_0(S)$ to the powerset $\mathcal{P}(\mathcal{C}_0(S'))$. To start, represent $\gamma \in \mathcal{C}_0(S)$ by a curve minimally intersecting $\partial S'$.

First suppose S' is not an annulus. If $i(\gamma, \partial S') = 0$, then either $\gamma \subset S'$, and we let $\pi_{S'}(\gamma) = \{\gamma\}$, or γ misses S', and we let $\pi_{S'}(\gamma)$ be the empty set. Otherwise, γ

intersects $\partial S'$. For each arc α of $\gamma \cap S'$, take the boundary of a regular neighborhood of $\alpha \cup \partial S'$ and exclude the component curves which are not essential in S'. Because S' is neither pants nor annulus, some curves remain. Let these comprise $\pi_{S'}(\gamma)$.

In the case that S' is an annulus, when $i(\gamma, \partial S') = 0$ we let $\pi_{S'}(\gamma)$ be the empty set. When γ intersects $\partial S'$, one expects $\pi_{S'}(\gamma)$ to consist of the arcs of γ intersecting S'. However, ambiguity arises because curves such as γ and $\partial S'$ are defined up to ∂S-fixing isotopy, but vertices of $\mathcal{C}(S')$ represent arcs up to $\partial S'$-fixing isotopy. To remedy this, give S a hyperbolic metric. Consider the cover of S corresponding to the fundamental group of S' embedded in that of S. This cover is a hyperbolic annulus endowed with a canonical "boundary at infinity" coming from the boundary of two-dimensional hyperbolic space (the unit circle, if one uses the Poincaré disk model). Name the closed annulus A, and let $\mathcal{C}(A)$ stand in for $\mathcal{C}(S')$. Each $\gamma \in \mathcal{C}_0(S)$ has a geodesic representative which lifts to A, and if γ intersects S', some of the lifts connect the boundary components. Let these comprise $\pi'_S(\gamma)$.

We will frequently say something like γ *projects to* S' to mean $\pi_{S'}(\gamma)$ is not the empty set. Where X is some collection of curves $\{\gamma_i\}$, let $\pi_{S'}(X) = \bigcup_i \pi_{S'}(\gamma_i)$. If $\pi_{S'}(X)$ is not empty, let $\text{diam}_{S'}(X)$ denote its diameter in $\mathcal{C}(S')$; omit the subscript S' to mean diameter of X in $\mathcal{C}(S)$ itself. We are content with a map from $\mathcal{C}(S)$ to subsets of $\mathcal{C}(S')$, rather than $\mathcal{C}(S')$ directly, precisely because multicurves have bounded-diameter projection:

Lemma II.1. *Suppose γ is a multicurve of S, and S' a subsurface. If γ projects to S', $\text{diam}_{S'}(\pi_{S'}(\gamma)) \leq 2$.*

This fact appears as Lemma 2.3 in [28]. Note that complexity ξ in [28] differs from our definition by three. Also, there is a minor error in Lemma 2.3, corrected in [30] (pg. 28), which we avoid by stating the lemma for multicurves rather than

simplices of $\mathcal{C}(S)$—these do not coincide if $\xi(S) = 1$.

When multicurves α and β both project to S', define their *projection distance* $d_{S'}(\alpha, \beta)$ as $\operatorname{diam}_{S'}(\pi_{S'}(\alpha) \cup \pi_{S'}(\beta))$. This is a "distance" in that it satisfies the triangle inequality and symmetry, but it need not discern curves: Lemma II.1 implies that disjoint multicurves have a projection distance of at most two, when defined. In the other extreme, one easily finds examples of curves close in $\mathcal{C}(S)$ with large projection distance in $\mathcal{C}(S')$—a small illustration of the great wealth of structure that opens up when one considers not only $\mathcal{C}(S)$ but curve complexes of all subsurfaces (see, for example, [28]).

2.3   Cut-coded subsurfaces and domains

We now present an alternate definition of subsurface that ducks the nuisance of disconnected subsurfaces containing annular components parallel to the boundary of other components. These are the only subsurfaces capable of being mutually nested (via isotopy) yet not topologically equivalent. Our technical antidote may seem tedious, but as an upside, it translates our current notion of a subsurface into an object encoded unambiguously by curves, a recurrent theme of this paper. Moreover, the new viewpoint facilitates the next section's definition of active subsurfaces, based on Ivanov's work on $\operatorname{Mod}(S)$ subgroups. The efficient reader is welcome to skim the definition, taking note of Lemmas II.2 and II.3, and rely on Figure 2.1 for intuition.

A *cut-coded subsurface* of S consists of two pieces of information: (1) a multicurve γ and (2) a partition of the non-pants components of $S\backslash\gamma$ into two sets: excluded and included components.

Let us clarify what subsurface S' this data is meant to describe. Label the component curves of γ by $\gamma_1, \ldots \gamma_n$, and the included components $A_1, \ldots A_m$. The

multicurve γ contains each component of ∂A_i, but some γ_j may not be a boundary curve for any A_i. Let N_j be regular neighborhoods of those γ_j belonging to no ∂A_i. Then S' consists of the union of the A_i, seen as subsurfaces, and the implied annuli N_j. Ignoring multiplicity, $\partial S'$ and γ are the same multicurve.
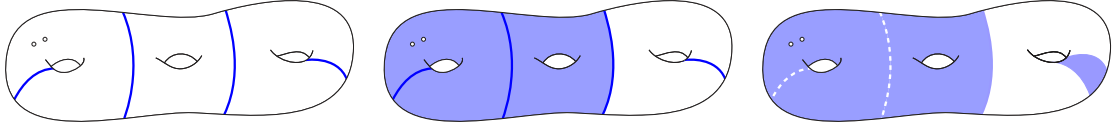

Figure 2.1: How to define a cut-coded subsurface

From left to right: (1) Set multicurve. (2) Choose included components (shaded). (3) Domains (shaded) correspond to included components and implied annuli.

Call the A_i *component domains* and the N_j *annular domains* of S'. We use *domain* to refer to either kind, or any subsurface that may appear as a domain—in other words, any connected (essential) non-pants subsurface.

Not all subsurfaces are cut-coded, as the latter never include pants, parallel annuli, or annuli parallel to the boundary of a component domain. Cut-coded subsurfaces are exactly those appearing as active subsurfaces of mapping classes, which the next section details.

The cut-coded subsurfaces of S admit a partial order $\subset$ detected by subsurface projection. Say a subsurface A *nests* in B if A may be isotoped into B. Now suppose A and B are cut-coded subsurfaces with domains A_i and B_j respectively. Say $A \subset B$ if every A_i nests in some B_j. Transitivity and reflexivity of $\subset$ are obvious, but antisymmetry is relatively special. If $A \subset B$ and $B \subset A$, one can check A and B are given by the same data as cut-coded subsurfaces, ultimately because they never contain an annulus parallel to another domain. In contrast, general disconnected subsurfaces can be mutually nested but not isotopic.

Call two subsurfaces *disjoint* if they may be isotoped apart, and *overlapping* if they are neither disjoint nor nested; note that overlapping subsurfaces are distinct by definition. Projection determines relations between domains:

Lemma II.2. *Suppose A and B are domains in S.*

(i) $\pi_A(\partial B)$ is empty if and only if A nests in B or its complement.

(ii) A and B overlap if and only if ∂A projects to B and ∂B to A.

(iii) If ∂A projects to B but $\pi_A(\partial B)$ is empty, A nests in B.

(iv) If $\pi_A(\partial B)$ and $\pi_B(\partial A)$ are both empty, A and B are equal or disjoint.

Proof. Statements (ii) – (iv) derive from (i). The forward implication of (i) requires connectedness of A: one may choose curves α_1 and α_2 that fill A, so that A is represented by a regular neighborhood of $\alpha_1 \cup \alpha_2$ with disks and annuli added to fill in homotopically trivial or ∂S-parallel boundary components. If $\pi_A(\partial B)$ is empty, then ∂B is disjoint from $\alpha_1 \cup \alpha_2$. Therefore the two curves, and consequently A itself, can be isotoped either entirely inside B or into its complement. The reverse implication is self-evident. $\square$

Consider a strictly increasing sequence $A^1 \subsetneq A^2 \subsetneq \cdots \subsetneq A^m$ of cut-coded subsurfaces of S. Let α be a maximal multicurve (i.e., a pants decomposition) including ∂A_i for all i. Each step of the sequence corresponds to some curve of α appearing for the first time as a component of ∂A_k or an essential curve in A_k, so the maximum length of the sequence is twice the number of components of α. We have just observed:

Lemma II.3. *If $\xi(S) \geq 1$, a strictly increasing sequence of nonempty cut-coded subsurfaces of S has length at most 2ξ.*

It is easy to construct sequences realizing the upper bound.

2.4 Active subsurfaces and the Omnibus Theorem

Thurston originally classified elements of $\text{Mod}(S)$ by their action on the space of projective measured foliations, a piecewise linear space obtained by completing and projectivizing the space of weighted curves on S. He defined a *pseudo-Anosov diffeomorphism* as one that fixes a pair of transverse projective measured foliations, and proved that any pseudo-Anosov mapping class (i.e., any mapping class fixing no multicurve) has a representative pseudo-Anosov diffeomorphism. In fact he proved more:

Theorem II.4 (Thurston, [36],[34]). *Every element f of $\text{Mod}(S)$ has a representative diffeomorphism F such that, after cutting S along some 1-dimensional submanifold C, F restricts to a pseudo-Anosov diffeomorphism on the union of some components of $S\backslash C$, and has finite-order on the union of the rest.*

Birman-Lubotzky-McCarthy proved that C represents a unique isotopy class, and they gave a simple way to find it [4]. We call this isotopy class the *canonical reduction multicurve* of f. Ivanov generalized to subgroups both the classification and the canonical cutting method [19]. We build the notion of an *active subsurface* according to this last, largest perspective. Let us emphasize that, while we aim to paint a picture that seems perfectly natural, the validity of the definitions in this section depends on multiple lemmas and theorems from [19], which serves as reference for assertions presented without proof.

Ivanov identified a congenial property of many mapping classes. Call a mapping class *pure* if, in the theorem above, it restricts to either the identity or a pseudo-Anosov on each component (in particular, no components are permuted). Because

one may take the empty set for C in Theorem II.4, any pseudo-Anosov mapping classes is pure. The property of being pure is most useful for reducible mapping classes, because cutting along C "reduces" S to a collection of smaller subsurfaces. The point in what follows is to formalize this procedure.

Call a subgroup pure if it consists entirely of pure mapping classes. Nontrivial pure mapping classes have infinite order, so they let us ignore the complications of torsion. Fortunately, the mapping class group has finite-index pure subgroups (Theorem 3, [19]). In particular, one can take the kernel of the homomorphism $\mathrm{Mod}(S) \twoheadrightarrow \mathrm{Aut}(H_1(S, Z/3\mathbb{Z}))$, induced by the action of $\mathrm{Mod}(S)$ on homology. Name this subgroup $\Gamma(S)$.

Let us first define the canonical reduction multicurve $\sigma(G)$ and active subsurface $\mathcal{A}(G)$ of a pure subgroup $G < \mathrm{Mod}(S)$. The multicurve $\sigma(G)$ consists of all curves γ such that (i) G fixes γ, and (ii) if some curve β intersects γ, then G does not fix β. The active subsurface $\mathcal{A}(G)$ is cut-coded with multicurve γ. Included components correspond to those on which some element of G "acts pseudo-Anosov." Specifically, for each component Q of $S \backslash \sigma(G)$, we have a homomorphism $\rho_Q : G \to \mathrm{Mod}(Q)$ such that $\rho_Q(g)$ is the mapping class of $F|_Q$, where F is a homeomorphism representing g. Let G_Q denote the image $\rho_Q(G)$. Because G is pure, for each component Q, the image G_Q either contains a pseudo-Anosov, in which case Q is included, or is the trivial group, in which case Q is excluded (Theorem 7.16, [19]).

It follows that the annular domains of $\mathcal{A}(G)$ correspond to neighborhoods of those $\gamma \in \sigma(G)$ that only bound components Q for which G_Q is trivial. Because by definition γ is not superfluous, some $g \in G$ restricts, in a neighborhood of γ, to a power of a Dehn twist about γ. Section 2.5 justifies why we consider Dehn twists annulus pseudo-Anosovs.

If G is an arbitrary subgroup, choose a finite-index pure subgroup G' and define $\sigma(G) = \sigma(G')$ and $\mathcal{A}(G) = \mathcal{A}(G')$. Any choice of G' gives the same multicurve, and one can always take $G' = G \cap \Gamma(S)$. G acts on $S \backslash \sigma(G)$, although its elements may permute the components. For each component Q, one can define ρ_Q on the finite-index subgroup of G stabilizing Q. Each image G_Q is finite or contains a pseudo-Anosov, and $\sigma(G)$ is the minimal multicurve with this property (Theorem 7.16, [19]). It follows that an infinite subgroup G contains a pseudo-Anosov if and only if $\sigma(G)$ is empty; let us call such a subgroup *irreducible*.

For any mapping class g, let $\sigma(g) = \sigma(\langle g \rangle)$ and $\mathcal{A}(g) = \mathcal{A}(\langle g \rangle)$. In this terminology, we recast Ivanov's Theorem 6.3 [19] as follows:

Theorem II.5. *For any $G < \mathrm{Mod}(S)$ there exists $f \in G$ such that $\mathcal{A}(f) = \mathcal{A}(G)$.*

After Lemma II.7 below, we can recognize Theorem II.5 as Mosher's Omnibus Subgroup Theorem. We do not require this theorem for our proofs, only the validity of the definitions on which it is based. In fact, one can prove Theorem I.1 with no mention of active subsurfaces for non-pure subgroups. However, its definition allows us to state the Main Theorem, Theorem III.1, which takes the more general, perhaps more useful point of view of Theorem II.5, adding the benefit of an f with bounded word length.

Facts about active subsurfaces occupy the remainder of this section. Directly from definitions, we derive Lemma II.6 below. With this we obtain Lemmas II.7 – II.9, which enable our proof of the main theorem. Let us refer to domains of $\mathcal{A}(G)$ or $\mathcal{A}(g)$ as domains of G or g respectively. Say G *moves* the curve $\gamma \in \mathcal{C}_0(S)$ if some element of G does not fix γ.

Lemma II.6. *Suppose H and G are pure subgroups of $\mathrm{Mod}(S)$ and $\gamma \in \mathcal{C}_0(S)$.*

(a) H moves γ if and only if γ projects to some domain of H.

(b) H and G fix and move the same curves if and only if $\mathcal{A}(H) = \mathcal{A}(G)$.

Lemma II.7. *If $H < G$ then $\mathcal{A}(H) \subset \mathcal{A}(G)$.*

Proof. We may replace H and G by the pure subgroups obtained by intersecting with $\Gamma(S)$. Suppose A is a domain of G and B a domain of H. Because G, and thus H, fixes ∂A, $\pi_B(\partial A)$ is empty. Lemma II.2 guarantees that B nests in A or its complement. Now suppose B is in the complement of every domain of G. Then any curve essential in B is fixed by G, thus by H. This contradicts that B is a domain of H, unless B has no essential curves. Thus B is an annulus and ∂B consists of (two copies of) a single curve β. Any γ intersecting β is moved by H, hence by G. But because G fixes γ, this means $\gamma \in \sigma(G)$. Because B is the annulus around γ, B isotopes into $\mathcal{A}(G)$. Thus every domain of H nests in $\mathcal{A}(G)$, which means $\mathcal{A}(H) \subset \mathcal{A}(G)$. $\qquad\square$
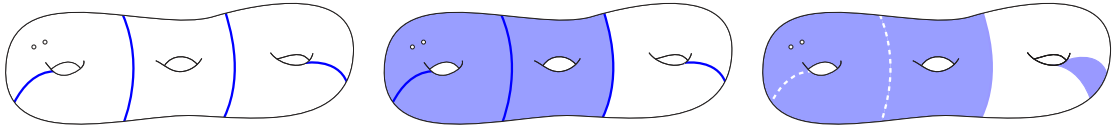

Figure 2.2: How to define the active subsurface of a group G

(1) Cut along canonical reduction multicurve. (2) Include components (shaded) on which G induces an irreducible subgroup. (3) Every element of G has a power supported on included components and implied annuli, by Lemma II.7. By Theorem II.5, some element has exactly this support.

In general, let $\mathcal{A}(X_1, X_2, \ldots)$ denote the active subsurface of the group generated by $X_1, X_2, \ldots$, where X_i may be either elements or subgroups of a mapping class group.

Lemma II.8. *Suppose H and G are pure elements or subgroups of $\mathrm{Mod}(S)$. Then $\mathcal{A}(H) \subset \mathcal{A}(G)$ if and only if $\mathcal{A}(G) = \mathcal{A}(G, H)$.*

Proof. Observe that $\mathcal{A}(H) \subset \mathcal{A}(G)$ implies that any curve fixed by G is fixed by H, and hence G and $\langle G, H \rangle$ fix the same curves. Apply Lemma II.6 to $G < \langle G, H \rangle$ for the forward direction, and Lemma II.7 to $H < \langle G, H \rangle$ for the reverse. □

Lemma II.9. *Let G be a pure subgroup of $\mathrm{Mod}(S)$ generated by $\{H, H_1, \ldots H_k, \ldots\}$, where H_k are groups or elements. Either $\mathcal{A}(H) = \mathcal{A}(G)$ or $\mathcal{A}(H) \neq \mathcal{A}(H, H_k)$ for some k.*

Proof. Suppose $\mathcal{A}(H) = \mathcal{A}(H, H_k)$ for all k. Let $K_i = \langle H, H_1, \ldots H_i \rangle$. Using induction and Lemma II.8 one can show $\mathcal{A}(H) = \mathcal{A}(K_i)$ for all i. Because the K_i exhaust G, Lemma II.7 implies $\mathcal{A}(g) \subset \mathcal{A}(H)$ for all $g \in G$. By Lemma II.6 one knows any curve fixed by H has empty projection to all $\mathcal{A}(g)$, so H and G fix and move the same curves, and consequently $\mathcal{A}(H) = \mathcal{A}(G)$. □

2.5 Machinations in the curve complex

This section collects several important curve complex results that provide the foundation for our proofs. Together, these results link mapping class behavior with curve complex geometry. Following our theme of interpreting $\mathrm{Mod}(S)$ via $\mathcal{C}(S)$, we note that these results also lead to a $\mathcal{C}(S)$-centric $\mathrm{Mod}(S)$ classification.

By definition, reducible mapping classes and finite-order mapping classes have bounded orbits in $\mathcal{C}(S)$. That pseudo-Anosovs have infinite-diameter orbits is corollary to a theorem of Masur and Minsky:

Theorem II.10 (Minimal translation of pseudo-Anosovs [27]). *There exists $c = c(S) > 0$ such that, for any pseudo-Anosov $g \in \mathrm{Mod}(S)$, vertex $\gamma \in \mathcal{C}_0(S)$, and nonzero integer n,*

$$d_S(g^n(\gamma), \gamma) \geq c|n|.$$

This gives us a way to recognize pseudo-Anosovs. It also suggests an alternate classification of elements on $\mathrm{Mod}(S)$, by whether they have finite, finite-diameter, or infinite-diameter orbits in $\mathcal{C}(S)$. If one takes that classification as a starting point, Dehn twists clearly qualify as pseudo-Anosovs for the annulus mapping class group.

For pure mapping classes, one may even refine this classification. Suppose $g \in \mathrm{Mod}(S)$ is pure and has a domain Y. The observation that, for all $\gamma \in \mathcal{C}_0(S)$, $\pi_Y(g(\gamma)) = g(\pi_Y(\gamma))$, yields:

Corollary II.11 (Minimal translation on subsurfaces). *There exists $c = c(S) > 0$ such that, for any pure element $g \in \mathrm{Mod}(S)$ with domain Y, vertex $\gamma \in \mathcal{C}_0(S)$ such that $\pi_Y(\gamma) \neq \emptyset$, and nonzero integer n,*

$$d_Y(g^n(\gamma), \gamma) \geq c|n|.$$

Thus every nontrivial pure element g has infinite-diameter orbits in the curve complexes of its domains. Moreover, an orbit of g in the curve complex of one of its domains projects to a bounded diameter subset of the curve complex of any subsurface properly nested in that domain. This one can prove using another theorem of Masur and Minsky, which along with Corollary II.11 is crucial to our core construction in Section 3.3:

Theorem II.12 (Bounded geodesic image [28]). *Let Y be a proper domain of S. Let $\mathcal{G}$ be a geodesic in $\mathcal{C}(S)$ whose vertices each project to Y. Then there is a constant $M = M(S)$ such that*

$$\mathrm{diam}_Y(\mathcal{G}) \leq M.$$

This theorem enabled Behrstock to obtain Lemma II.13 below, which Chapter III frequently employs. Here we give the elementary proof, with constructive constants,

communicated by Chris Leininger. Most of it appeared previously in [26]; this version adds the possibility of annular domains.

Lemma II.13 (Behrstock [2]). *For any pair of overlapping domains Y and Z and any multicurve x projecting to both,*

$$\mathrm{d}_Y(x, \partial Z) \geq 10 \implies \mathrm{d}_Z(x, \partial Y) \leq 4$$

Proof. First we gather the facts that prove the lemma when neither Y nor Z is an annulus. Suppose S' is a subsurface of S and $\xi(S), \xi(S') \geq 1$. Let u_0 and v_0 be curves on S which minimally intersect S' in sets of arcs. Suppose a_u is one these arcs for u_0, and u a component of the boundary of a neighborhood of $a_u \cup \partial S'$; suppose a_v and v play the same role for v_0. Then $u \in \pi_{S'}(u_0)$ and $v \in \pi_{S'}(v_0)$. Define intersection number of arcs to be minimal over isotopy fixing the boundary setwise but not necessarily pointwise. One has:

(1) If $i(a_u, a_v) = 0$, then $\mathrm{d}_{S'}(u_0, v_0) \leq 4$

(2) If $i(u, v) > 0$, then $i(u, v) \geq 2^{(\mathrm{d}_{S'}(u,v)-2)/2}$

(3) $i(u, v) \leq 2 + 4 \cdot i(a_u, a_v)$

Statement (1) follows from the proof of Lemma II.1 (Lemma 2.3 in [28]). Straightforward induction proves (2), which Hempel records as Lemma 2.1 in [16]. Fact (3) is the observation that essential curves from the regular neighborhoods of $a_u \cup \partial S'$ and $a_v \cup \partial S'$ intersect at most four times near every intersection of a_u and a_v, plus at most two more times near $\partial S'$.

Now assume $\xi(Y), \xi(Z) \geq 1$. Because $\mathrm{d}_Y(x, \partial Z) \geq 10 > 2$, the diameter is realized by curves $u \in \pi_Y(x), v \in \pi_Y(\partial Z)$ such that, by (2), $i(u, v) \geq 2^4 = 16$. By the definition of π_Y, these u and v come from arcs $a_u \subset x \cap Y$ and $a_v \subset \partial Z \cap Y$

respectively. By (3), $i(a_u, a_v) \geq (16 - 2)/4 > 3$. Thus a_u is an arc of x intersected thrice by an arc a_v of ∂Z, within the subsurface Y. Observe that one of the segments of a_u between points of intersection must lie within Z. This segment is an arc a_x of x disjoint from arcs of ∂Y in Z. Fact (1) implies $\mathrm{d}_Z(x, \partial Y) \leq 4$.

The main idea of this proof works for annular domains after a few more relevant facts. Endow S with a hyperbolic metric and let A be an embedded annulus with geodesic core curve α; let $\widetilde{A}$ be the corresponding annular cover of S. Let u and v be geodesic curves in S with lifts $\widetilde{u}$ and $\widetilde{v}$ traversing the core curve of $\widetilde{A}$. We have already mentioned

(4) $\mathrm{d}_A(u, v) = i(\widetilde{u}, \widetilde{v}) + 1$.

Let $\widetilde{\alpha}$ be the unique lift of α corresponding to the core curve of $\widetilde{A}$, and let $\widetilde{\alpha}_1$ and $\widetilde{\alpha}_2$ be the first lifts of α intersecting $\widetilde{u}$ on each side. On the sides of $\widetilde{\alpha}_i$ opposite $\widetilde{\alpha}$, $\widetilde{A}$ is isometric to its pre-image in the universal cover $\widetilde{S} = \mathbb{H}^2$. Because geodesics intersect only once in $\mathbb{H}^2$,

(5) at most two of the intersections of $\widetilde{u}$ and $\widetilde{v}$ occur outside the open segment of $\widetilde{u}$

between the $\widetilde{\alpha}_i$.

Now we can retrace the proof above, augmenting it to address the possibility that Y or Z is an annulus. Under some hyperbolic metric, $x, \partial Y$, and ∂Z have geodesic representatives. Suppose Y is an annulus with geodesic core curve y and $\widetilde{Y}$ is the corresponding annular cover of S. If $\mathrm{d}_Y(x, \partial Z) \geq 10$, (4) implies that some lifts $\widetilde{x}$ and $\widetilde{\partial Z}$ in $\widetilde{Y}$ intersect at least nine times. By (5), at least three of these intersections occur on an open segment of $\widetilde{x}$ between consecutive lifts of y, and a neighborhood of this segment embeds in S. As before, one finds an arc of x intersecting ∂Z at least three times in a neighborhood disjoint from ∂Y. If Z is not an annulus, one repeats

the conclusion of the first argument.

On the other hand, suppose that Z is an annulus with geodesic core curve z, and Y is any domain. The arguments thus far tell us one has an arc of x thrice intersecting z in a neighborhood of that arc disjoint from ∂Y. In the annular cover corresponding to Z, the three intersections correspond to a lift $\widetilde{x}$ of x intersecting the closed lift $\widetilde{z}$ of z and adjacent lifts $\widetilde{z}_1$, $\widetilde{z}_2$ on each side. Any lift of ∂Y cannot intersect $\widetilde{x}$ between these intersections, so (5) implies $i(\widetilde{x}, \widetilde{\partial Y}) \leq 2$. Fact (4) implies $d_Z(x, \partial y) \leq 3$. $\qquad \square$
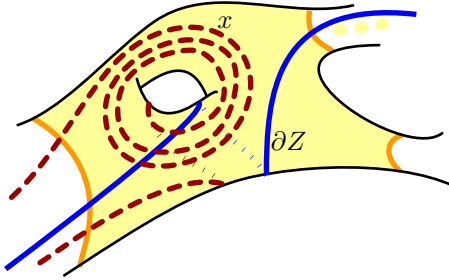

Figure 2.3: Key point of Behrstock's lemma

Because x and ∂Z have large projection distance in Y, one can find an arc of x intersecting an arc of ∂Z three times. Extra-fine dashed lines represent where curves run behind the surface.
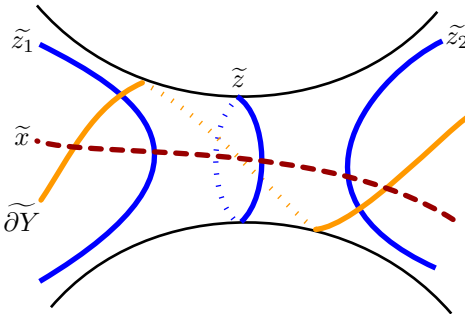

Figure 2.4: Behrstock's lemma for the annulus

If x and z intersect three times with no intersections of ∂Y in between, then in the annular cover corresponding to z, $\widetilde{x}$ can intersect $\widetilde{\partial Y}$ no more than twice. Extra-fine dashed lines represent parts of curves on the far side of the annulus.

Remark. Behrstock's lemma implies that $\mathcal{C}(S)$-orbits of a mapping class g have bounded projection to subsurfaces that overlap with domains of g. Using stronger results, Theorem II.12 in particular, one can prove that the g-orbit of any $\gamma \in \mathcal{C}(S)$ projects to an unbounded set in the curve complex of S' if and only if the orbit projects to S' and S' is a domain of g. Thus one obtains a refined, curve-complex-based classification of elements of $\mathrm{Mod}(S)$, by letting the phrase "g is pseudo-Anosov on S'" mean that the g-orbit of some $\gamma \in \mathcal{C}(S)$ has infinite-diameter projection to $\mathcal{C}(S')$.

CHAPTER III

The Recipe

We begin at the end, proving the Main Theorem, modulo two propositions, in Section 3.1. There, we reduce the proof to the problem of generating a short-word full-support mapping class when the generating set has only two pure elements; even then, we must simultaneously accommodate multiple scenarios happening on different, non-interacting subsurfaces. The actual construction of pseudo-Anosovs is left to Section 3.3, where in fact we identify all pseudo-Anosov elements in any group generated by two "sufficiently different" pure reducible mapping classes. That is, we prove Proposition I.2 of the introduction, which gives more than is needed for the main theorems. In general one is not so lucky as to start out with the condition on generators required for Proposition I.2 to apply. Thus we also need a recipe for writing a sufficiently different pair of mapping classes, given an arbitrary pair of pure reducible mapping classes generating an irreducible subgroup. Proposition III.4 fills that need; its proof occupies Section 3.2.

3.1   Writing the short word

Restated in the terminology introduced in Section 2.4, our goal is to prove the following:

Theorem III.1 (Main Theorem). *There exists a constant $K = K(S)$ with the property that, for any subset $\Sigma \subset \text{Mod}(S)$, there exists $f \in \langle \Sigma \rangle$ with Σ-length less than K, such that $\mathcal{A}(f) = \mathcal{A}(\Sigma)$.*

First, let us narrow our starting point. By definition, the finite-index pure subgroup $H = \langle \Sigma \rangle \cap \Gamma(S)$ has the same active subsurface as $\langle \Sigma \rangle$. Suppose the index of H in $\langle \Sigma \rangle$ is d. Lemma 3.4 of Shalen and Wagreich [35] provides a generating set for H of words less than $2d - 1$ in length according to the original generating set. Although they state the lemma for finite generating sets, nothing prevents the proof from applying to a general group. Therefore, if we find a full-support mapping class in H whose word length is less than l_1 in the new generating set, its word length is less than $l_1(2d-1)$ in the original generating set. Recall that $\Gamma(S)$ is the kernel of the $\text{Mod}(S)$ action on homology with coefficients in $\mathbb{Z}/3\mathbb{Z}$; its index, $|\text{Aut}(H_1(S, \mathbb{Z}/3\mathbb{Z}))|$, gives an upper bound for d.

Let $\Sigma' = \{h_i\}$ be the new generating set for H. Renumbering as necessary, Lemmas II.3 and II.9 provide a sequence $\mathcal{A}(h_1) \subsetneq \mathcal{A}(h_1, h_2) \subsetneq \mathcal{A}(h_1, h_2, h_3) \dots$ terminating in at most $2\xi(S)$ steps at $\mathcal{A}(h_1, h_2, \dots, h_k) = \mathcal{A}(G)$. Let $p_1 = h_1$ and suppose for each subgroup $\langle p_{i-1}, h_i \rangle$ we can spell a full-support element p_i (i.e. $\mathcal{A}(p_i) = \mathcal{A}(p_{i-1}, h_i)$) with word length less than l_2 in the generating set $\{p_{i-1}, h_i\}$. By Lemma II.7, $\mathcal{A}(h_i) \subset \mathcal{A}(p_i)$; inductively assuming that $\mathcal{A}(p_{i-1}) = \mathcal{A}(h_1, \dots, h_{i-1})$, we also know $\mathcal{A}(h_j) \subset \mathcal{A}(p_{i-1}) \subset \mathcal{A}(p_i)$ for all $j < i$. Then Lemmas II.7 – II.9 tell us $\mathcal{A}(p_i) = \mathcal{A}(h_1, \dots, h_i)$, for all i. In particular, p_k is full-support for H and has Σ'-length less than l_2^k, where $k \leq 2\xi$.

We have reduced Theorem III.1 to the case where A consists of a pair of pure mapping classes a and b. Let $G = \langle a, b \rangle$. Finding an element of full-support for G is equivalent to finding a pseudo-Anosov on $\mathcal{A}(G)$, although we contend with the fact

that $\mathcal{A}(G)$ may well be disconnected. On every domain of $\mathcal{A}(G)$, one of the following possibilities occurs:

(i) a and b are pseudo-Anosov.

(ii) a and b are both reducible.

(iii) One of a and b is pseudo-Anosov and the other is reducible.

To handle the first possibility, we quote a theorem of Fujiwara. An earlier incarnation of this theorem inspired the short-word question in the first place. Call a pair of pseudo-Anosovs *independent* if all pairs of nontrivial powers fail to commute. In torsion-free groups such as those we consider, either two pseudo-Anosovs generate a cyclic subgroup, or they are independent (see the proof of Theorem 5.12 in [19]). In the latter case, Fujiwara's theorem applies:

Theorem III.2 (Fujiwara [10], partial version)**.** *There exists a constant $L = L(S)$ with the following property. Suppose $a, b \in \mathrm{Mod}(S)$ are independent pseudo-Anosovs. Then for any $n, m \geq L$, $\langle a^n, b^m \rangle$ is an all-pseudo-Anosov rank-two free group.*

Case (ii) relies on two results. The first, Proposition III.3 below, is a subset of Proposition I.2 from the introduction; both are proved in Section 3.3. Define $Q = Q(S) = \max\{3, (2M + 4)/c\}$ where $M = M(S)$ is the constant from Theorem II.12 and $c = c(S)$ is the minimal translation length from Corollary II.11. Recall that a pair of pure reducible mapping classes a and b are *sufficiently different* if some $A \in \mathcal{A}(a)$ and $B \in \mathcal{A}(b)$ together fill S.

Proposition III.3. *For any $m, n > Q$ and sufficiently different pure mapping classes a and b, the group $\langle a^m, b^n \rangle$ is a rank-two free group and its elements are each either pseudo-Anosov or conjugate to a power of a or b.*

Proposition III.3 would be irrelevant if not for Proposition III.4, whose proof we leave to Section 3.2. Let $\xi = \xi(S)$ and $c = c(S)$ as above. The following proposition works for any domain $S' \subset S$.

Proposition III.4. *Suppose a and b are pure reducible mapping classes in $\mathrm{Mod}(S')$ and $\langle a, b \rangle$ is irreducible. For any $n \geq \xi - 1$ and $k \geq 20/c$, the words*

$$a_1 = (b^k a^k)^n b^k \cdot a \cdot ((b^k a^k)^n b^k)^{-1} \quad and \quad b_1 = (a^k b^k)^n a^k \cdot b \cdot ((a^k b^k)^n a^k)^{-1}$$

are sufficiently different pure reducible mapping classes.

We handle case (iii) by converting it to either case (i) or (ii), so with the three results above we may finish the proof. Recall we face the situation where a and b are pure reducible mapping classes generating a group G with possibly disconnected active subsurface. Our task is to write a word in a and b that induces a pseudo-Anosov on every domain of G.

Let a_1 and b_1 be as in Proposition III.4. Because these are simply conjugates of a and b, they fulfill the same possibilities (i)-(iii) on each domain of $\mathcal{A}(G)$. Let $\bar{L} = \max\{L(S') : S' \in T\}$, where T is the finite set of topological types of domains of S (i.e., connected non-pants subsurfaces), and L is the constant in Theorem III.2. Similarly let $\bar{Q} = \max\{Q(S') : S' \in T\}$. Choose $P \geq \max\{\bar{L}, \bar{Q}\}$. Consider the following word:

$$w = b_1^P a_1^P b_1^{-P} a_1^P$$

On domains where possibility (i) holds, either a_1 and b_1 commute and $w = a_1^{2P}$, a pseudo-Anosov, or a_1 and b_1 are independent and w is pseudo-Anosov by Theorem III.2. On domains where possibility (ii) holds, w is pseudo-Anosov by Propositions III.3 and III.4. On domains where possibility (iii) holds, and a_1 is the pseudo-Anosov, we see by re-writing w as $b_1^P a_1^P b_1^{-P} \cdot a_1^P$ that it is the product of powers of

pseudo-Anosovs, so we may proceed as we did for case (i). Otherwise a_1 is reducible, and for any $\gamma \in \sigma(a_1)$, $\mathrm{d}_X(b_1^P(\gamma), \gamma) \geq Pc \geq Qc \geq 3$, where the leftmost inequality employs Corollary II.11 and the rest are by construction. In particular, $\gamma \in \sigma(a_1)$ and $b_1^P(\gamma) \in \sigma(b_1^P a_1 b_1^{-P})$ together fill the domain, so a_1 and $b_1^P a_1 b_1^{-P}$ are sufficiently different pure reducible mapping classes. Then Proposition III.3 guarantees w is pseudo-Anosov. Thus on all domains of $\mathcal{A}(G)$, w is pseudo-Anosov, meaning w is full-support for G.

In terms of $\{a, b\}$ the word length of w is $4P \cdot (2k(2n+1)+1)$. Therefore one can let $l_2 = 4P \cdot (80\xi/c + 1)$. For Theorem III.1 one may take

$$K(S) = 2 \cdot |\mathrm{Aut}(H_1(S, \mathbb{Z}/3\mathbb{Z}))| \cdot (320P\xi/c + 4P)^{2\xi}.$$

$\square$

3.2 Finding sufficiently different reducibles

The purpose of this section is to prove Proposition III.4. In the hypothesis of the proposition, a and b are pure reducible elements of $\mathrm{Mod}(S')$, where S' is a domain in S, and $\langle a, b \rangle$ is irreducible. Given any $k \geq 20/c(S)$ and $n \geq \xi(S) - 1$, let $u = (b^k a^k)^n b^k$ and $v = (a^k b^k)^n a^k$. To prove the proposition, we show that any choice of $\alpha \in \sigma(uau^{-1})$ and $\beta \in \sigma(vbv^{-1})$ together fill S'. Necessarily α and β will bound some domains $A \in \mathcal{A}(a)$ and $B \in \mathcal{A}(b)$, and A and B together fill S'.

Without loss of generality we may assume $S' = S$.

We prove Proposition III.4 in two steps. First, we use Behrstock's Lemma (II.13) to follow the image of any $\alpha_1 \in \sigma(a)$ and $\beta_1 \in \sigma(b)$ under the composition of alternations of high powers of a and b. We find that $u(\alpha_1)$ and $v(\beta_1)$ have a certain subsurface projection property. The second step is to prove that this property implies the two curves fill S.

For the purpose of step one, we introduce the *overlap graph* for a pair of pure mapping classes. Suppose a and b are pure nontrivial mapping classes generating the group G. Let $\{A_i\}_{1 \leq i \leq m}$ and $\{B_j\}_{1 \leq j \leq n}$ be the domains of a and b respectively. The overlap graph $\mathcal{O}(a, b)$ consists of a vertex for every domain of a that overlaps with some domain of b, and for every domain of b that overlaps with some domain of a—recall that two domains overlap if the boundary of each domain projects to the other. Edges connect domains that overlap, so that one may color the vertices depending on whether they represent domains of a or b, obtaining a bipartite graph. Assign each edge length one so that $\mathcal{O}(a, b)$ has the usual path metric.
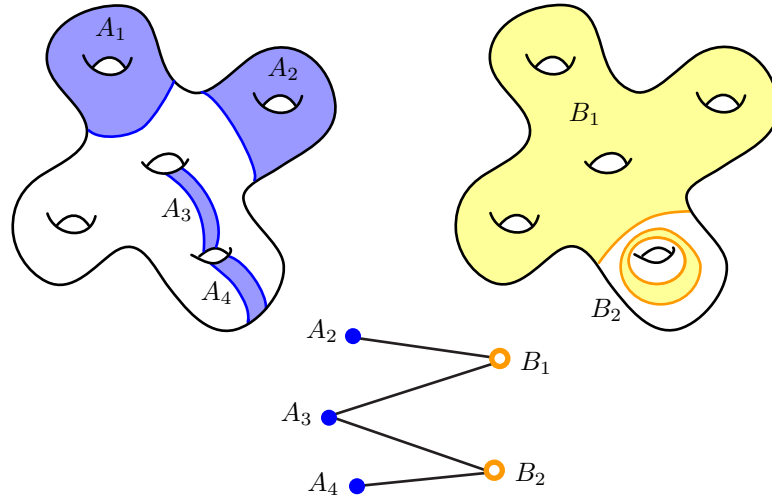
Figure 3.1: Example of an overlap graph.

Note that the foregoing definition works whether or not G contains a pseudo-Anosov. Now let us assume that, as in the hypotheses for Proposition III.4, G does contain a pseudo-Anosov, or equivalently, $\mathcal{A}(G)$ is the connected surface S'. In particular a and b fix no common curve. If a is not pseudo-Anosov, then any domain A_i has essential boundary consisting of curves not fixed by b. This boundary must project to some domain B_j of b. If B_j has no essential boundary, then it is

all of S' and b is pseudo-Anosov. Otherwise, the boundary projects to some domain of a. If ∂B_j projects to A_i, then B_j and A_i overlap. If not, then A_i nests in B_j by Lemma II.2, and ∂B_j projects to some other domain A_k. In this case, ∂A_k in turn projects to B_j, because otherwise B_j nests in A_k, implying A_i nests in A_k, a contradiction. So B_j and A_k overlap. In either case $\mathcal{O}(a, b)$ has at least two vertices and an edge connecting them. These observations imply

Lemma III.5. *If $\mathcal{A}(G)$ is connected and $\mathcal{O}(a, b)$ is empty, then at least one of a and b is pseudo-Anosov.*

In the case where $\mathcal{O}(a, b)$ is not empty, we distinguish between the domains represented by vertices—call these domains *overlappers*—and those not represented. Revisiting the discussion above, we can extract a fact so handy we should name it:

Lemma III.6. *When neither a nor b are pseudo-Anosov, any non-overlapper of one nests in an overlapper of the other.*

That observation helps us to the result below.

Lemma III.7. *If $\mathcal{A}(G)$ is connected, then $\mathcal{O}(a, b)$ is connected.*

Proof. Assuming $\mathcal{O}(a, b)$ is not empty, we show how to find a path between any two vertices. Realize the domains as closed submanifolds with minimal pairwise intersection (i.e., ensure that boundary curves intersect essentially and use pairwise disjoint submanifolds to represent domains for the same mapping class). For any two vertices X and Y of $\mathcal{O}(a, b)$, choose points x and y in the corresponding domains, and connect these by a path p. Letting $A = \bigcup_i A_i$ and $B = \bigcup_j B_j$, one sees that each component of $S' \backslash (A \cup B)$ is a disk or boundary-parallel annulus with a boundary component consisting of pieces of $\partial A \cup \partial B$. Thus one can isotope p to lie entirely within $A \cup B$, and furthermore transversal to the boundary curves. The path also lies

entirely within overlappers, by Lemma III.6. Tracing the path gives a finite sequence of domains, each overlapping with the neighbor before and after, alternating between A_i's and B_j's by design. The same sequence appears as a path in $\mathcal{O}(a, b)$ connecting X and Y. □

Now we can state precisely the first step to proving Proposition III.4. Assume a and b are as in Proposition III.4, with the same definitions as above for $A = \bigcup_i A_i$ and $B = \bigcup_j B_j$. Recall $c = c(S)$ is the minimal translation constant from Theorem II.11, and $\xi(S)$ is surface complexity.

Lemma III.8. *Suppose α_1 and β_1 are curves in $\sigma(a)$ and $\sigma(b)$ respectively, and k and n are integers such that $k \geq 20/c(S)$, and $n \geq \xi(S) - 1$. For $\alpha = (b^k a^k)^n b^k(\alpha_1)$ and $\beta = (a^k b^k)^n a^k(\beta_1)$, the following hold for any choice of A_i or B_j.*

(i) If ∂A projects to B_j, then so does α, and $d_{B_j}(\alpha, \partial A) \geq 14$.

(ii) If ∂B projects to A_i, then so does β, and $d_{A_i}(\beta, \partial B) \geq 14$.

Proof. By symmetry we need only prove (i), the case for α. We can use the overlap graph to track the image of α_1 under alternating applications of b^k and a^k. For ease of exposition, we will not always distinguish between overlapper domains and their representative vertices in $\mathcal{O}(a, b)$.

Let Y_0 be the vertices of $\mathcal{O}(a, b)$ corresponding to the overlappers of b that intersect α_1; note that Y_0 cannot be empty. Let X_0 be the vertices adjacent to Y_0. For $i \in \mathbb{N}$, let Y_i be the vertices adjacent to X_{i-1}, and let X_i be the vertices adjacent to Y_i. In other words, Y_i corresponds to b-domains that overlap with a-domains of X_{i-1}, and X_i are the a-domains that overlap with b-domains of Y_i.

Observe that the vertices of Y_{i+1} lie within a radius-two neighborhood of Y_i. Because $\xi(S)$, the maximum number of disjoint curves on S, gives an upper bound on

the number of domains for a mapping class in $\mathrm{Mod}(S)$, twice $\xi(S)$ bounds the number of vertices in $\mathcal{O}(a,b)$. Because $\mathcal{O}(a,b)$ is connected, one sees that for $n \geq \xi(S) - 1$, Y_n consists of all the b-vertices, corresponding to all the overlapper domains of b. If a domain B_j of b is not an overlapper, then by Lemma III.6 it nests in some domain of a, which precludes projection of ∂A to B_j. To prove Lemma III.8 it suffices to establish the following claim:

(∗) For any domain Y of Y_n, $\mathrm{d}_Y((b^k a^k)^n b^k(\alpha_1), \partial A) \geq 14$

We induct on n. By definition, α_1 projects to any domain Y of Y_0. Applying Theorem II.11,

$$\mathrm{d}_Y(b^k(\alpha_1), \alpha_1) \geq c|k| \geq 20 \geq 14$$

Because α_1 is a component of the multicurve ∂A, and because projection distances are diameters, we obtain the $n = 0$ case of (∗):

$$\mathrm{d}_Y(b^k(\alpha_1), \partial A) \geq 14$$

Supposing (∗) true for $n = m$, we prove it for $n = m + 1$. Consider any domain Y of Y_m, and recall that $\mathrm{diam}_Y(\partial A) \leq 2$, by Lemma II.1. For any A_i overlapping Y, the triangle inequality gives

$$\mathrm{d}_Y((b^k a^k)^m b^k(\alpha_1), \partial A_i) \geq 12$$

In particular, $(b^k a^k)^m b^k(\alpha_1)$ projects to every A_i that overlaps some Y of Y_m; those A_i are exactly the vertices of X_m. In other words, for any $X \in X_m$, there exists $Y \in Y_m$ such that

$$\mathrm{d}_Y((b^k a^k)^m b^k(\alpha_1), \partial X) \geq 12$$

We may apply Behrstock's lemma, giving

$$\mathrm{d}_X((b^k a^k)^m b^k(\alpha_1), \partial Y) \leq 4$$

On the other hand, Theorem II.11 guarantees

$$d_X(a^k(b^ka^k)^mb^k(\alpha_1), (b^ka^k)^mb^k(\alpha_1)) \geq c|k| \geq 20$$

Employing the triangle inequality,

$$d_X(a^k(b^ka^k)^mb^k(\alpha_1), \partial Y) \geq 16$$

The above holds for any $Y \in Y_m$, and $X \in X_m$. Given any $Y' \in Y_{m+1}$, recall that Y' corresponds to a domain overlapping some $X \in X_m$, which in turn overlaps with some $Y \in Y_m$. We also know $\text{diam}_X(\partial Y \cup \partial Y') \leq 2$, because $\partial Y'$ and ∂Y are disjoint. Another triangle inequality gives

$$d_X(a^k(b^ka^k)^mb^k(\alpha_1), \partial Y') \geq 14$$

Again we can apply Behrstock's lemma. In fact we can mirror the last four inequalites:

$$d_{Y'}(a^k(b^ka^k)^mb^k(\alpha_1), \partial X) \leq 4$$

$$d_{Y'}(b^ka^k(b^ka^k)^mb^k(\alpha_1), a^k(b^ka^k)^mb^k(\alpha_1)) \geq 20$$

$$d_{Y'}(b^ka^k(b^ka^k)^mb^k(\alpha_1), \partial X) \geq 16$$

$$d_{Y'}(b^ka^k(b^ka^k)^mb^k(\alpha_1), \partial X') \geq 14$$

where X' is any domain of X_{m+1}. Note that $\partial X'$ is a subset of ∂A, and recall Y' was an arbitrary vertex of Y_{m+1}. Thus we may re-write the last inequality to give the $m + 1$ step of the induction claim: for any $Y \in Y_{m+1}$,

$$d_Y((b^ka^k)^{m+1}b^k(\alpha_1), \partial A) \geq 14$$

$\square$

Remark. The lower bound for k in Lemma III.8 is slightly higher than necessary in order to simplify statements and proofs for both Lemmas III.8 and III.9.

We are ready for the second step to proving Proposition III.4. Its own proof will probably feel like déjà vu. Again, a and b are as in the proposition, with the same definitions as above for $A = \bigcup_i A_i$ and $B = \bigcup_j B_j$. Note that in the following lemma, α and β are arbitrary.

Lemma III.9. *Suppose $\alpha, \beta \in \mathcal{C}_0(S')$ satisfy the following:*

(i) If ∂A projects to B_j, then so does α, and $\mathrm{d}_{B_j}(\alpha, \partial A) \geq 14$.

(ii) If ∂B projects to A_i, then so does β, and $\mathrm{d}_{A_i}(\beta, \partial B) \geq 14$.

Then α and β fill S'.

Proof. Given an arbitrary curve $\gamma \in \mathcal{C}(S')$, we show it intersects either α or β. Because γ cannot be fixed by both a and b, it projects to some domain, and we can choose this domain to be an overlapper (Lemma III.6). Without loss of generality we may assume γ projects to a domain Y of b, which overlaps with a domain X of a. Because multicurves have diameter-two projections (Lemma II.1), properties (i) and (ii) imply:

$$\mathrm{d}_Y(\alpha, \partial X) \geq 12$$

$$\mathrm{d}_X(\beta, \partial Y) \geq 12$$

Assuming γ does not intersect α, we show it must intersect β. Again Lemma II.1 implies $\mathrm{d}_Y(\alpha, \gamma) \leq 2$. The following inequalities employ the triangle inequality, Behrstock's lemma, and another triangle inequality.

$$\mathrm{d}_Y(\gamma, \partial X) \geq 10$$

$$\mathrm{d}_X(\gamma, \partial Y) \leq 4$$

$$\mathrm{d}_X(\gamma, \beta) \geq 8$$

In yet another instance of Lemma II.1, the last inequality shows γ and β must intersect. $\qquad\square$

Proof of Proposition III.4. Let $u = (b^k a^k)^n a^k$ and $v = (a^k b^k)^n b^k$. Taken together, Lemmas III.8 and III.9 show that, for any $\alpha \in \sigma(uau^{-1})$ and $\beta \in \sigma(vbv^{-1})$, α and β fill S'. Therefore $a_1 = uau^{-1}$ and $b_1 = vbv^{-1}$ are sufficiently different pure reducible mapping classes. $\qquad\square$

3.3 Constructions

Here we give the pseudo-Anosov construction at the heart of the proof of our Main Theorem. That is, we prove Proposition III.3 of Section 3.1, which states that, in a group generated by two sufficiently different pure reducible mapping classes, every element is pseudo-Anosov except those conjugate to powers of the generators. The second part of this section upgrades Proposition III.3 to Proposition I.2 given in the introduction, by proving the all-pseudo-Anosov subgroups are convex cocompact.

3.3.1 Of pseudo-Anosovs

Recall that a pseudo-Anosov is characterized by having infinite-diameter orbits in the curve complex. Thus we will know w is a pseudo-Anosov if distances $\mathrm{d}(\gamma, w^n(\gamma))$ grow as n increases, where γ is some vertex in $\mathcal{C}(S)$. Proposition III.3 is an application of Lemma III.10 below, which uses the Bounded Geodesic Image Theorem (Theorem II.12) to glean geometric information about sequences of geodesics. The author learned this "bootstrapping" strategy from Chris Leininger, who takes a similar tack in [25]. In the same vein one also has Proposition 5.2 in [22], but here the domains are not restricted.

Let $M = M(S)$ be the constant from Theorem II.12. Below, properties (ii) and (iii) ensure that the projection in (iv) is well-defined.

Lemma III.10. *Suppose $\{Y_j\}$ is a sequence of domains in S, and $\{X_j\}$ a sequence of subsets of $\mathcal{C}_0(S)$, for which the following properties hold for all j:*

(i) $\operatorname{diam}(X_j) \leq 2$

(ii) X_j *and* X_{j+1} *are pairwise disjoint*

(iii) *The set of curves that project trivially to* Y_j *is a subset of* X_j

(iv) $d_{Y_j}(w_{j-1}, w_{j+1}) > 2M$ *for any choice of* $w_{j-1} \in X_{j-1}$, $w_{j+1} \in X_{j+1}$

Then for any $w_i \in X_i$ and $w_{i+k} \in X_{i+k}$, the geodesic $[w_i, w_{i+k}]$ contains a vertex from X_j for $i \leq j \leq i+k$. Also, the X_j are pairwise disjoint. In particular, $[w_i, w_{i+k}]$ has length at least k.

Proof. First let us induct on k the claim that, for any $w_i \in X_i$ and $w_{i+k} \in X_{i+k}$, the geodesic $[w_i, w_{i+k}]$ contains a vertex from X_j for $i \leq j \leq i+k$. This is vacuously true for $k = 1$.

The general induction step works even for $k = 2$, but let us separate this case to highlight its use of the Bounded Geodesic Image Theorem, Theorem II.12. By property (iv), $d_{Y_{i+1}}(w_i, w_{i+2}) > 2M > M$, so Theorem II.12 requires $[w_i, w_{i+2}]$ contain a vertex v disjoint from Y_{i+1}. By property (iii), v is contained in X_{i+1}. Of course, the endpoints of $[w_i, w_{i+2}]$ lie in X_i and X_{i+2}, so the induction claim holds.

Now assume $k > 2$ and for any $w_i \in X_i, w_{i+k-1} \in X_{i+k-1}$, the geodesic $[w_i, w_{i+k-1}]$ intersects X_j for $i \leq j \leq i+k-1$. The main task is to show $[w_i, w_{i+k}]$ intersects X_{i+k-1}.

Choosing any $w_{i+k-2} \in X_{i+k-2}$, our first step is to show one can pick a geodesic $[w_i, w_{i+k-2}]$ avoiding X_{i+k-1}. Suppose we are given a geodesic that doesn't, that is,

for some $w'_{i+k-1} \in X_{i+k-1}$,

$$[w_i, w_{i+k-2}] = [w_i, w'_{i+k-1}] \cup [w'_{i+k-1}, w_{i+k-2}].$$

Let us require w'_{i+k-1} to be the first vertex along $[w_i, w_{i+k-2}]$ belonging to X_{i+k-1}, so that only the last vertex of $[w_i, w'_{i+k-1}]$ lies in X_{i+k-1}. The induction hypothesis applied to the segment $[w_i, w'_{i+k-1}]$ implies that, for some $w'_{i+k-2} \in X_{i+k-2}$,

$$[w_i, w_{i+k-2}] = [w_i, w'_{i+k-2}] \cup [w'_{i+k-2}, w'_{i+k-1}] \cup [w'_{i+k-1}, w_{i+k-2}].$$

Property (ii) ensures the second and third segments above each have length at least one. Thus their union $[w'_{i+k-2}, w_{i+k-2}]$ has length at least two. Property (i) tells us we can replace it with a length-2 geodesic contained entirely in X_{i+k-2}, giving a new $[w_i, w_{i+k-2}]$ avoiding X_{i+k-1}.

Property (iii) ensures that every vertex of $[w_i, w_{i+k-2}]$ projects nontrivially to Y_{i+k-1}, as does w_{i+k}. Projection distances make sense and Theorem II.12, the Bounded Geodesic Image Theorem, applies:

$$\mathrm{d}_{Y_{i+k-1}}(w_i, w_{i+k-2}) \leq \mathrm{diam}_{Y_{i+k-1}}([w_i, w_{i+k-2}]) \leq M.$$

Finally, a triangle inequality:

$$
\begin{aligned}
\mathrm{diam}_{Y_{i+k-1}}([w_i, w_{i+k}]) \;\geq\;&\; \mathrm{d}_{Y_{i+k-1}}(w_i, w_{i+k}) \\
\geq\;&\; \mathrm{d}_{Y_{i+k-1}}(w_{i+k-2}, w_{i+k}) - \mathrm{d}_{Y_{i+k-1}}(w_i, w_{i+k-2}) \\
>\;&\; 2M - M = M
\end{aligned}
$$

Again by Theorem II.12, we know $[w_i, w_{i+k}]$ intersects X_{i+k-1}. The end of the induction is easy: for some w_{i+k-1} in X_{i+k-1},

$$[w_i, w_{i+k}] = [w_i, w_{i+k-1}] \cup [w_{i+k-1}, w_{i+k}].$$

The induction hypothesis says the first segment on the right intersects each X_j, $i \leq j \leq i+k-1$. So the geodesic on the left intersects X_j, $i \leq j \leq i+k$, as required.

Finally we check that the sets X_j are pairwise disjoint. Suppose $z \in X_i \cap X_{i+k}$ for some nonzero k. By the part of the lemma already proved, the geodesic $[z, z]$ contains vertices in X_j for $i \leq j \leq i+k$: simply put, all those X_j intersect at z. But consecutive X_j do not intersect. $\qquad\square$

Proof of Proposition III.3. Recall that $Q = Q(S) = \max\{3, (2M+4)/c\}$ where $M = M(S)$, the constant from Theorem II.12, and $c = c(S)$, the constant from Corollary II.11. The hypothesis states that a and b are pure reducible mapping classes, with $A \in \mathcal{A}(a)$ and $B \in \mathcal{A}(b)$, and A and B together fill S. We will show that for any $m, n > Q$, $\langle a^m, b^n \rangle$ is free and nontrivial elements of $\langle a^m, b^n \rangle$ are either pseudo-Anosov or conjugate to powers of a^m or b^n.

If S is a torus or four-punctured sphere, the only pure reducible mapping classes are Dehn twists about curves, and any pair of curves fill the surface. In this case the result for $Q = 3$ is due to [13]; see also [36, 33, 17]. For the rest of the proof let us assume $\xi(S) > 1$.

Let $n, m \geq Q$ be arbitrary integers. Let α be a component of ∂A and β of ∂B. Let $C_A \subset \mathcal{C}_0(S)$ be the vertices with empty projection to A, and define C_B analogously. Observe that C_A and C_B sit in 1-neighborhoods of α and β, respectively, so they each have diameter 2 in $\mathcal{C}(S)$. Because any curve intersects either A or B, C_A and C_B contain no common vertices—this is the only place we use the fact that A and B fill S.

Here's the whole point of our choice of Q: for any nonzero integer k,

$$\mathrm{d}_A(C_B, a^{mk}(C_B)) \geq \mathrm{d}_A(\beta, a^{mk}(\beta)) \geq |mk|c > Qc \geq 2M + 4$$

$$d_B(C_A, b^{nk}(C_A)) \geq d_B(\alpha, b^{nk}(\alpha)) \geq |nk|c > Qc \geq 2M + 4$$

Notice that the domains A, B and sets C_A, C_B play the same roles for a^m and b^n as they do for a and b. To ease the burden of excessive exponents, we replace a^m with a and b^n with b for the remainder of the proof. In this new notation, our goal is to show that $\langle a, b \rangle$ is free, and every nontrivial element of $\langle a, b \rangle$ is pseudo-Anosov except those conjugate to powers of a or b. We have already established that, for any nonzero k,

$$d_A(C_B, a^k(C_B)) > 2M + 4 \qquad \text{and} \qquad d_B(C_A, b^k(C_A)) > 2M + 4$$

In particular, for any $\gamma, \gamma' \in C_B$ or $\delta, \delta' \in C_A$,

$$(3.1) \qquad d_A(\gamma, a^k(\gamma')) > 2M \qquad \text{and} \qquad d_B(\delta, b^k(\delta')) > 2M$$

We use this shortly to apply Lemma III.10.

In the abstract free group on a and b, a word w has reduced form $w = s_1 s_2 \ldots s_R$, where the *syllables* s_i are nontrivial powers of either a or b, and s_i is a power of a if and only if $s_{i\pm1}$ is a power of b (i.e. powers of a and b alternate). Define *powerblind* word length $|\cdot|_*$ by $|w|_* = R$, the number of syllables of w.

Claim. For any word w, either $d(w(\alpha), \alpha) \geq |w|_*$ or $d(w(\beta), \beta) \geq |w|_*$.

It immediately follows that $\langle a, b \rangle$ is a rank-two free group. If $|w|_*$ is even, it is easy to check $|w^n|_* = n|w|_*$. The claim implies $d(w^n(\gamma), \gamma) \geq n|w|_*$, where γ is α or β, depending on w. In particular the orbit of w has infinite diameter in the curve complex, so w is a pseudo-Anosov. If $|w|_*$ is odd and neither conjugate to a power of a nor of b, then it is conjugate to v such that $|v|_*$ is even. As v is pseudo-Anosov, so is its conjugate w.

It remains to prove the claim. Towards this, we describe a sequence of domains and $\mathcal{C}_0(S)$-subsets fulfilling the hypotheses of Lemma III.10. Let w_0 be the identity,

$w_1 = s_1$, $w_2 = s_1 s_2$, and so forth, so that w_i is the word formed by the first i syllables of w. If s_1 is a power of a, let $I(a)$ correspond to the even integers, and $I(b)$ to the odds; if s_1 is a power of b, switch the roles of even and odd. Define R-length sequences of vertices γ_r, domains Y_r, and sets X_r as follows.

$$\gamma_r = w_r(\alpha) \qquad Y_r = w_r(A) \qquad X_r = w_r(C_A) \qquad \forall r \in I(a) \cap [0, R]$$

$$\gamma_r = w_r(\beta) \qquad Y_r = w_r(B) \qquad X_r = w_r(C_B) \qquad \forall r \in I(b) \cap [0, R]$$

In addition, set $\{\gamma_{-1}, Y_{-1}, X_{-1}\}$ equal to $\{\beta, B, C_B\}$, if s_1 is a power of a, or $\{\alpha, A, C_A\}$, if s_1 is a power of b.

Each X_j is isometric to C_A or C_B, and furthermore pairs X_j, X_{j+1} are isometric to the pair C_A, C_B, disregarding order. Therefore the sequence $\{Y_j, X_j\}$ meets conditions (i) – (iii) of Lemma III.10. Condition (iv) requires $\mathrm{d}_{Y_j}(v_{j-1}, v_{j+1}) > 2M$ for any choice of $v_{j-1} \in X_{j-1}$, $v_{j+1} \in X_{j+1}$. For $j = 0$, this condition simply restates one of the inequalities in (3.1) above, so let us suppose $j \geq 1$. Without loss of generality, assume $Y_j = w_j(A)$ and $v_{j-1} = w_{j-1}(\gamma_-)$, $v_{j+1} = w_{j+1}(\gamma_+)$ for some $\gamma_-, \gamma_+ \in C_B$. Because subsurface projection commutes naturally with the action of the mapping class group,

$$\mathrm{d}_{w_j(A)}(v_{j-1}, v_{j+1}) = \mathrm{d}_A(w_j^{-1}(v_{j-1}), w_j^{-1}(v_{j+1})) = \mathrm{d}_A(s_j^{-1}(\gamma_-), s_{j+1}(\gamma_+))$$

Exactly one of s_j and s_{j+1} is a power of a, while the other is a power of b. In any case one knows that, for some $\gamma, \gamma' \in C_B$ and nonzero k,

$$\mathrm{d}_A(s_j^{-1}(\gamma_-), s_{j+1}(\gamma_+)) = \mathrm{d}_A(a^k(\gamma), \gamma')$$

We used the fact that b fixes C_B setwise. Applying inequality (3.1) above, we can conclude $\mathrm{d}_{w_j(A)}(v_{j-1}, v_{j+1}) > 2M$.

By Lemma III.10, the geodesics $[\gamma_{-1}, \gamma_R]$ and $[\gamma_0, \gamma_R]$ have lengths at least $R + 1$ and R respectively. Depending on w, one of these geodesics is either $[\alpha, w(\alpha)]$ or $[\beta, w(\beta)]$. This proves the claim, thus the lemma. $\square$
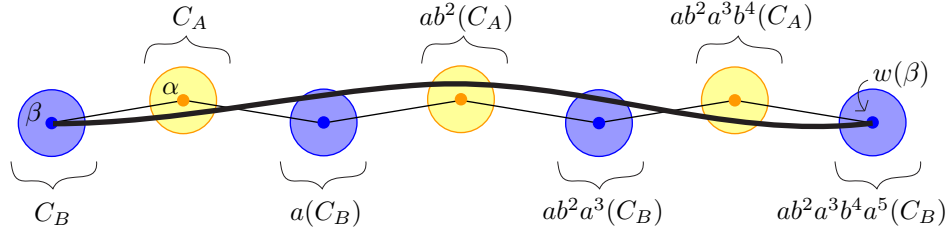

Figure 3.2: The sequence defined in the proof of Proposition III.3

Here $w = ab^2a^3b^4a^5$. Vertices correspond to γ_i and shaded circles represent X_i, which each sit in the 1-neighborhood of γ_i. The heavy, smooth line is a geodesic connecting β and $w(\beta)$; it intersects each of the X_i.

Remark. One can ask whether Proposition III.3 above can be proven for n-tuples of pure mapping classes subject to appropriate conditions. For example, if pairwise they fulfill the requirements for a and b in the lemma, do sufficiently high powers generate rank n free groups? One would need to avoid the trivial counterexamples arising from redundant generating sets such as $\{a, b, a^k b a^{-k}\}$. Excluding this latter possibility, the immediate adaptation of the proof of Lemma III.3 for n-tuples only gives a Q that depends on the particular n-tuple; proving that some Q works for any n-tuple seems to require more maneuvering.

3.3.2   Of convex cocompact all-pseudo-Anosov subgroups

Now we upgrade Proposition III.3 to Proposition I.2, stated in the introduction. Proposition I.2 gives new examples of mapping class subgroups which are convex cocompact. The notion of convex cocompactness for mapping class groups was introduced by Farb and Mosher in [8]. Later Kent and Leininger, and separately Hamenstädt, proved that $G < \mathrm{Mod}(S)$ is convex cocompact if and only if its orbit

embedding in the curve complex is a quasi-isometry ([21], [12]), that is, for some $\gamma \in \mathcal{C}_0(S)$, $K > 1$, and $C \geq 0$, and for all $w \in G$,

$$(3.2) \qquad\qquad K|w| + C \geq \mathrm{d}(\gamma, w(\gamma)) \geq |w|/K - C$$

where $|w|$ gives word length with respect to some metric on G. It is proved in [8], and immediate after [21] and [12], that non-trivial elements of a convex cocompact group are pseudo-Anosov. Convex cocompact subgroups which are free of finite rank are called Schottky. Farb and Mosher proved these are common in mapping class groups of closed surfaces, and Kent and Leininger include a new proof which also works for non-closed surfaces:

Theorem III.11 (Abundance of Schottky groups [8], [21])**.** *Given a finite set of pseudo-Anosovs $\{g_1, g_2, ...g_k\}$ which are independent (i.e., no pairs of powers commute), there exists l such that for all $m > l$, $\{g_1^m, g_2^m, ...g_k^m\}$ is Schottky.*

Fujiwara found a uniform bound for the above theorem in the case of two-generator subgroups:

Theorem III.12 (Fujiwara [10], full version)**.** *There exists a constant $L = L(S)$ with the following property. Suppose $a, b \in \mathrm{Mod}(S)$ are independent pseudo-Anosov elements. Then for any $n, m \geq L$, $\langle a^n, b^m \rangle$ is Schottky.*

Proposition I.2 provides a source of Schottky subgroups with arbitrary rank. It simply adds to Proposition III.3 the claim that finitely generated all-pseudo-Anosov subgroups of $\langle a, b \rangle$ are Schottky. Recall that a and b are pure mapping classes with essential reduction curves α and β respectively, such that $\alpha \cup \beta$ fills S. Proposition III.3 tells us that $G = \langle a, b \rangle$ is a free group, and its all-pseudo-Anosov subgroups are exactly those containing no conjugates of powers of a or b. For example, any subgroup of the commutator group $[G, G]$ qualifies.

Proof of Proposition I.2. Let $G = \langle a, b \rangle$ as in the proposition, and let H be a finitely generated, all-pseudo-Anosov subgroup of G. To show that H is Schottky, we need quasi-isometry constants for the inequalities in (3.2). For an orbit embedding of a finitely generated group, the upper bound of (3.2) always comes for free, using any K greater than the largest distance some generator translates γ. So our work is the lower bound. In the proof of Lemma III.3 we saw that $\mathrm{d}(\gamma, w(\gamma)) \geq |w|_*$ where γ is one of α or β, depending on w. It is not hard to check that, in general, $\mathrm{d}(\alpha, w(\alpha)) \geq |w|_* - 1$ (the same is true replacing α with β). However, powerblind word length is not a word metric with respect to any finite generating set for H. In what follows, we define a convenient generating set for H such that, letting $|\cdot|_H$ denote the corresponding word metric, we have $K|w|_* \geq |w|_H$, where K is the size of this generating set. Then we can conclude that $\mathrm{d}(\alpha, w(\alpha)) \geq |w|_H/K - 1$, completing the proof.

The existence of this convenient generating set has no relation to our setting of subgroups of mapping class groups, so we isolate this fact as a separate technical lemma. We only need that G is a rank two free group generated by a and b. Suppose H is a finitely generated subgroup. For $w \in H$, let $|w|$ denote its length in G with respect to the generating set $\{a, b\}$. Let $H_l = \{h \in H \colon |h| \leq l\}$ and choose L so that H_L generates H. Define $|\cdot|_H$ as word length in H with respect to H_L. Let K be the number of elements in H_L.

Lemma III.13. *If H contains no element conjugate to a power of a or b, then* $K|w|_* \geq |w|_H.$

Proof. Suppose $w \in H$ has length $|w|_H = n$, and $w = h_1 h_2 \cdots h_n$, where $h_i \in H_L$. One knows that $|h_i h_{i+1}| > |h_i|, |h_{i+1}|$ because otherwise $h_i h_{i+1} \in H_L$, which contradicts that $|w|_H = n$ (one could take a shorter path to w in the Cayley graph

of H with respect to H_L, by replacing h_i and h_{i+1} with their product, another generator). In particular this means that strictly less than half of the word h_i, written as a product of a's and b's, gets canceled by a piece of the word h_{i+1} in a's and b's. Likewise, strictly less than half gets canceled by a piece of the word h_{i-1}. Therefore, at least the middle letter of h_i, if $|h_i|$ is odd (the middle pair of letters if $|h_i|$ is even) gets contributed to the $\{a, b\}$-spelling of the word w. Incidentally, this is showing that $|w| \geq |w|_H$, implying that finitely generated subgroups of the rank-two free group are quasi-isometrically embedded.

Call the middle letter or pair of letters of each h_i its *core*. Powerblind word length $|w|_*$ can be shorter than $|w|_H$ only if a string of consecutive h_i's, say, $h_i h_{i+1} \cdots h_k$, all have a or a^2 at their core, or if they all have b or b^2 at their core, and these cores contribute to the same syllable (power of a or b) in the $\{a, b\}$-spelling of w. In that case one can write, for $i \leq j \leq k$,

$$h_j = u_j \cdot x^{e(j)} \cdot v_j$$

where x is a or b and $x^{e(j)}$ includes the core of each h_j. Furthermore, $v_j = u_{j+1}^{-1}$, so that

$$h_i h_{i+1} \cdots h_k = u_i x^N v_k$$

where $N = \sum_{i \leq j \leq k} e(j)$. It is possible u_i or v_k are empty words, but N cannot be zero, because otherwise $|h_i h_{i+1} \cdots h_k| \leq |h_i|/2 + |h_k|/2 \leq L$, meaning the entire string can be replaced with a single element of H_L, contradicting the fact that $|w|_H = n$. In this context, suppose $h_i = h_k$. Then $u_k = u_i x^p$ for some p. But because $v_{k-1} = u_k^{-1}$, one has

$$h_i \cdots h_{k-1} = u_i x^{N'} u_i^{-1}$$

for $N' = N - e(k) - p$. However, the stipulation that H contains no elements of G

conjugate to powers of a or b, precludes this scenario.

To summarize, if any consecutive string $h_i \cdots h_k$ in the H_L-spelling of w contributes to the same syllable in the $\{a, b\}$-spelling of w, that string includes at most one instance of each element of H_L. We have demonstrated a correspondence between letters h_i and syllables of w written with respect to H_L and $\{a, b\}$ respectively: each letter corresponds to at least one syllable (the one in which its core appears), and at most K letters correspond to the same syllable. Therefore $K|w|_* \geq |w|_H$. $\quad \square$

As described above, Lemma III.13 completes the proof of Proposition I.2. $\quad \square$

CHAPTER IV

And More

4.1 A mapping class group "strong" Tits alternative

It is not known whether mapping class groups are linear, except in special cases: most notably, Krammer and Bigelow proved linearity of braid groups (mapping class groups of punctured spheres) [24], [3]. However, many linear group results inspire mapping class group analogues. For example, McCarthy and Ivanov independently proved that mapping class groups satisfy a version of the Tits alternative: either a subgroup is virtually abelian or it contains a rank-two free group [29], [19]. We gave a quantified version in the introduction, which we restate and prove below.

Theorem IV.1. *Suppose S is a compact orientable surface with mapping class group* $\mathrm{Mod}(S)$ *and $G < \mathrm{Mod}(S)$ is finitely generated and not virtually abelian. There exists a constant w depending only on S such that, for any finite set Σ generating G, some pair of elements with Σ-length less than w generates a rank two free group.*

Proof. By replacing S with $\mathcal{A}(G) \subset S$, we may assume G contains a pseudo-Anosov. Theorem I.1 guarantees a pseudo-Anosov f with Σ-length less than K. We may also replace G with its finite-index pure subgroup $H = G \cap \Gamma(S)$—by construction, f is already in H. Using Lemma 3.4 of [35], the proof of Theorem I.1 also provides a generating set Σ' for H consisting of words with Σ-length less than K. Suppose

$h \in \Sigma'$. Either hfh^{-1} and f are independent or, appealing to Corollary 3.6 in [19], we know h fixes the stable foliation of f in the space of projective measured foliations. If the latter option holds for all $h \in \Sigma'$, then Lemma 5.10 in [19] concludes that H is cyclic. But that implies G is virtually abelian, contradicting our hypotheses. Therefore for some $h \in \Sigma'$, hfh^{-1} and f are independent. Then Theorem III.2 by Fujiwara provides the constant L such that f^L and hf^Lh^{-1} generate a free group, and each have length less than $w = K(L+2)$. $\square$

Theorem IV.1 was first proved by the author in [26], with a simpler proof that does not require one to construct pseudo-Anosovs.

4.2 Uniform uniform exponential growth

Let G be a group generated by a finite set Σ. Denote by $b(G, \Sigma, n)$ the number of elements of G which may be written as words in Σ with length less than n. The exponential growth rate of G with respect to Σ is

$$h(G, \Sigma) = \limsup_{n \to \infty} \frac{\log(b(G, \Sigma, n))}{n}$$

and the minimal exponential growth rate of G is

$$h(G) = \inf\{h(G, \Sigma) : \Sigma \text{ generates } G\}.$$

G is said to have exponential growth if $h(G, \Sigma) > 0$ for some and hence all Σ, and uniform exponential growth if $h(G) > 0$.

Exponential growth does not generally imply uniform exponential growth [37], but Eskin, Mozes, and Oh proved this implication holds for linear groups over a field with characteristic zero [7]. Breuillard and Gelander strengthened the uniformity: for example, they proved that for a finitely generated, discrete subgroup $G < GL_n(\mathbb{R})$ which is not virtually nilpotent, $h(G)$ has a lower bound depending only on n [6].

Corollary IV.2 mirrors this property for mapping class groups of surfaces. In fact, the mapping class group has uniform exponential growth, because its action on homology yields a surjection onto a linear group with uniform exponential growth [1]. Corollary IV.2 states that all $\mathrm{Mod}(S)$ subgroups with exponential growth have uniform exponential growth, and furthermore their minimal growth rates have a lower bound determined by the surface:

Corollary IV.2. *The minimal growth rate $h(G)$ of any non-elementary subgroup of $G > \mathrm{Mod}(S)$ has a positive lower bound $r = r(S)$ depending only on the surface.*

Proof. By counting only the elements of the free group found in Theorem IV.1, one can calculate a lower bound on the growth rate. Specifically, $r = (\log 3)/w$, where w is the constant from Theorem IV.1 $\qquad\square$

4.3   Bounded return probability

Another application of Theorem IV.1 is motivated by a comment in [5]. The existence of short independent words yields a fact about simple random walks: return probability decays exponentially with a rate depending only on w and the initial support of the walk. In other words, start at the identity and iterate a uniform probability distribution on a symmetric set (i.e., one including elements and their inverses) Σ to obtain a simple random walk on the group $\langle\Sigma\rangle$. Let $p^{(k)}$ be the probability that one returns to the identity in k steps. The quantity $\rho = \lim_{k\to\infty}(p^{(k)})^{\frac{1}{k}}$ is the Kesten spectral radius, introduced in [23], and regarding this one can show the corollary below.

Corollary IV.3. *There exists a function $F : \mathbb{N} \to (0,1)$ depending only on S which fulfills the following: for the simple random walk described above where $\Sigma \subset \mathrm{Mod}(S)$, $|\Sigma| = n$, and $\langle\Sigma\rangle$ is not virtually abelian, one has $\rho \leq F(n)$.*

Proof. For a reference on random walks, see [38]; also, the author enjoys the exposition in [20]. For a symmetric set Σ, let $\rho(\Sigma)$ be the Kesten spectral radius for the corresponding simple random walk on $\langle\Sigma\rangle$. Part of Theorem 10.3 in [38], specialized to simple random walks, gives the following equivalence: $\rho(\Sigma)$ is strictly less than 1 if and only if there exists $\kappa(\Sigma) > 0$ such that, for all finitely supported functions $f : \langle\Sigma\rangle \to \mathbb{C}$,

$$\sum_{x \in \langle\Sigma\rangle} f(x)^2 \leq \kappa(\Sigma) \sum_{x \in \langle\Sigma\rangle, g \in \Sigma} (f(x) - f(xg))^2/2.$$

One may think of the sum on the right as the sum of squares of differences over edges in the Cayley graph determined by Σ. For the forward implication, one may choose $\kappa \geq (1 - \rho)^{-1}$; backwards, one knows $\rho \leq 1 - \kappa^{-1}$ (note this corrects a typo in the proof in [38]).

Suppose Σ is as in the hypothesis: $\Sigma \subset \mathrm{Mod}(S)$, $|\Sigma| = n$, and $\langle\Sigma\rangle$ is not virtually abelian. Theorem IV.2 gives two elements $u, v \in \langle\Sigma\rangle$ with Σ-length less than w such that u and v freely generate F_2. Let T be the symmetric set $\{u, v, u^{-1}, v^{-1}\}$. Then $\rho(T) = \sqrt{3}/2$, because this is the simple random walk on standard generators of F_2, already calculated in [23]. From $\rho(T)$ one obtains $\kappa(T)$, and, because cosets of $\langle T\rangle$ are for this purpose indistinguishable from $\langle T\rangle$,

$$\sum_{x \in \langle\Sigma\rangle} f(x)^2 \leq \kappa(T) \sum_{x \in \langle\Sigma\rangle, t \in T} (f(x) - f(xt))^2/2$$

for any finitely supported function $f : \langle\Sigma\rangle \to \mathbb{C}$. Furthermore, because the generators in T are of the form $g_1 \cdots g_d$, $g_i \in \Sigma$, $d \leq w$, the triangle inequality implies that for any $x \in \langle\Sigma\rangle, t \in T$, there exists $y \in \langle\Sigma\rangle, g \in \Sigma$ such that

$$(f(x) - f(xt))^2 \leq w^2(f(y) - f(yg))^2.$$

In the Cayley graph, the pair (y, yg) corresponds to an edge of the path connecting x to xt. It is straightforward to count that for a given $y \in \langle\Sigma\rangle$ and $g \in \Sigma$, the

difference $f(y) - f(yg)$ appears at most $w(n-1)^{w-1}$ times in the upper bound for some $f(x) - f(xt)$, $x \in \langle \Sigma \rangle, t \in T$. Therefore one has

$$\sum_{x \in \langle \Sigma \rangle} f(x)^2 \leq \kappa(T) w^3 (n-1)^{w-1} \sum_{x \in \langle \Sigma \rangle, g \in \Sigma} (f(x) - f(xg))^2 / 2.$$

Thus

$$\rho(\Sigma) \leq 1 - \frac{1 - \sqrt{3}/2}{w^3(n-1)^{w-1}}.$$

The right hand side defines $F(n)$. □

4.4 Slow-growing groups

Breuillard proved the precise analogue to Theorem IV.1 in the linear group setting—in fact, he coined the phrase "strong Tits alternative" [5]. That is, for G a finitely generated, non-virtually-solvable subgroup of $GL_n(K)$, one has the same conclusion as Theorem IV.1, with w depending only on n. One cannot hope for a bound on w independent of dimension, in the linear group setting, or surface complexity, in the mapping class group setting. This is because there exist linear groups with arbitrarily small minimal growth rates ([11]; see Remark 1.4 in [5]), and these same examples embed into mapping class groups, by a construction suggested by Breuillard. This gives the following theorem:

Theorem IV.4 (Breuillard-Mangahas, in [26]). *Let $w = w(S)$ and $r = r(S)$ be the constants from Theorem IV.1 and Corollary IV.2. There exists a sequence of surfaces S_n such that $w(S_n) \to \infty$ and $r(S_n) \to 0$ as $n \to \infty$.*

For the proof, it suffices to find subgroups of mapping class groups with arbitrarily small minimal growth rates. In [11], Grigorchuk and de la Harpe construct a sequence of groups G_n with the property that $h(G_n) \to 0$ as $n \to \infty$. Following a construction suggested by Emmanuel Breuillard, one may represent these groups in

mapping class groups. In fact, the method shows that finite extensions of mapping class subgroups are themselves mapping class subgroups. Let B be the non-empty boundary of a surface S, $\mathrm{Homeo}^+(S, B)$ denote the group of orientation-preserving homeomorphisms of S fixing B pointwise, and $\mathrm{Mod}(S, B)$ be $\mathrm{Homeo}^+(S, B)$ modulo equivalence by isotopy keeping B fixed throughout.

Proposition IV.5. *If a finitely generated group G has a finite index subgroup H isomorphic to a subgroup of $\mathrm{Mod}(S, B)$, then G embeds in the mapping class group of a connected surface.*

Proof. Throughout, we consider H as a subgroup of both $\mathrm{Mod}(S, B)$ and G. By passing to a finite index subgroup, we may assume H is normal in G. Ultimately we embed G in the mapping class group of a surface S^* modeled on a Cayley graph of G/H with a copy of S for each vertex, and annuli for edges.

Before constructing S^*, let us address two technical issues. First: we need a group of homeomorphisms associated to H, which will let us define S^* as a quotient. Let F_Σ be the free group on Σ, where $\Sigma = \{g_1, g_2, ...g_m\}$ is a finite generating set for G. Let $\bar{H}$ be the pre-image of H under the canonical map $r : F_\Sigma \to G$; observe $F_\Sigma/\bar{H}$ and G/H are isomorphic finite groups. As a subgroup of a free group, $\bar{H}$ is also a free group, and so the map $r : \bar{H} \to H$ lifts to $\tilde{r} : \bar{H} \to \mathrm{Homeo}^+(S, B)$. Second: without loss of generality we assume that (1) B has at least $2m$ components, and (2) only the identity in H fixes all curves on S (i.e., no element of H is a Dehn twist about components of B). We can ensure this by replacing S with a larger surface whose mapping class group also contains H [32].

Now we build S^*. To each generator g_i associate a pair of boundary components B_i^+ and B_i^-, and fix orientation-reversing homeomorphisms $b_i : B_i^+ \to B_i^-$. Let S' be the quotient of the disjoint union $F_\Sigma \times S$ by the equivalence $(k, x) \sim (kh^{-1}, \tilde{r}(h)(x))$

for all h in $\bar{H}$. One obtains a copy of S for each coset $g\bar{H}$, corresponding to each vertex of the Cayley graph of $F_\Sigma/\bar{H}$. Let B' be the image of $F_\Sigma \times B$ in S'. For a connected surface, obtain S^* by identifying components of B' as specified by the edges of the Cayley graph: $[(k, x)] \sim [(kg_i, b_i(x))]$ for all g_i and $x \in B_i^+$.

The action of F_Σ on $F_\Sigma \times S$ by $g \cdot (k, x) = (gk, x)$ descends to an action on S^* that gives a homomorphism $p : F_\Sigma \to \mathrm{Homeo}^+(S^*)$. Let q be the quotient $q : \mathrm{Homeo}^+(S^*) \to \mathrm{Mod}(S^*)$. Define a homomorphism $\phi : G \to \mathrm{Mod}(S^*)$ by $\phi(g) = qpr^{-1}(g)$. Although r is not injective, ϕ is well-defined if, for any two words w and v representing the same element in G, $qp(wv^{-1})$ is the identity e. Notice wv^{-1} represents the identity in G, so it is in $\bar{H}$, thus

$$p(wv^{-1}) \cdot [(k, x)] = [(wv^{-1}k, x)] = [(k, \tilde{r}(k^{-1}wv^{-1}k)(x))]$$

Because r maps $k^{-1}wv^{-1}k$ to the identity, the lift $\tilde{r}$ maps it to a homeomorphism isotopic to the identity on S, thus on components of S', and thus on S^* (because all homeomorphisms and isotopies fix B', the boundary identifications are no obstruction). Hence $p(wv^{-1})$ is isotopic to the identity, so $qp(wv^{-1}) = e$.

Finally, we show that ϕ is injective. Suppose $\phi(g) = e$ and choose $w \in r^{-1}(g)$. Because $qp(w) = e$, $p(w)$ is isotopic to the identity. In particular, $p(w)$ fixes the subsurface of S^* corresponding to the component $[\{e\} \times S]$ of S'. Therefore w represents an element of H. Furthermore, $p(w)$ fixes the isotopy classes of all curves on that subsurface. Then by condition (2) on H described above, w represents the identity, so $g = r(w) = e$. □

Proof of Theorem IV.4. As remarked above, we only need to show that the slow-growing groups G_n are subgroups of mapping class groups. Proposition IV.5 applies because each G_n has a finite-index subgroup H_n isomorphic to the direct product

of 2^n finitely generated free groups [11]. The H_n may be realized as mapping class subgroups using pairs of Dehn twists on disjoint subsurfaces of S_{2^n} minus 2^{n+1} disks: mapping classes supported on disjoint subsurfaces commute, and a pair of properly chosen Dehn twists generates the rank two free group, which contains all other finitely generated free groups. $\qquad\square$

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] J. W. Anderson, J. Aramayona, and K. J. Shackleton, *Uniformly exponential growth and mapping class groups of surfaces*, In the tradition of Ahlfors-Bers. IV, Contemp. Math., vol. 432, Amer. Math. Soc., Providence, RI, 2007, pp. 1–6.

[2] Jason Behrstock, *Asymptotic geometry of the mapping class group and Teichmüller space*, Geom. Topol. **10** (2006), 1523–1578.

[3] Stephen J. Bigelow, *Braid groups are linear*, J. Amer. Math. Soc. **14** (2001), no. 2, 471–486 (electronic).

[4] Joan S. Birman, Alex Lubotzky, and John McCarthy, *Abelian and solvable subgroups of the mapping class groups*, Duke Math. J. **50** (1983), no. 4, 1107–1120.

[5] E. Breuillard, *A strong Tits alternative*, preprint (2003), arXiv:0804.1395v1.

[6] E. Breuillard and T. Gelander, *A topological Tits alternative*, Ann. of Math. (2) **166** (2007), no. 2, 427–474.

[7] Alex Eskin, Shahar Mozes, and Hee Oh, *Uniform exponential growth for linear groups*, Int. Math. Res. Not. (2002), no. 31, 1675–1683.

[8] Benson Farb and Dan Margalit, *A primer on mapping class groups*, in preparation, http://www.math.utah.edu/~margalit/primer/.

[9] Koji Fujiwara, *Subgroups generated by two pseudo-Anosov elements in a mapping class group. I. Uniform exponential growth*, Groups of diffeomorphisms, Adv. Stud. Pure Math., vol. 52, Math. Soc. Japan, Tokyo, 2008, pp. 283–296.

[10] ———, *Subgroups generated by two pseudo-Anosov elements in a mapping class group. II. Uniform bound on exponents*, preprint (2009), arXiv:0908.0995v1.

[11] Rostislav Grigorchuk and Pierre de la Harpe, *Limit behaviour of exponential growth rates for finitely generated groups*, Essays on geometry and related topics, Vol. 1, 2, Monogr. Enseign. Math., vol. 38, Enseignement Math., Geneva, 2001, pp. 351–370.

[12] Ursula Hamenstädt, *Word hyperbolic extensions of surface groups*, preprint (2005), arXiv:0807.4891v2.

[13] Hessam Hamidi-Tehrani, *Groups generated by positive multi-twists and the fake lantern problem*, Algebr. Geom. Topol. **2** (2002), 1155–1178 (electronic).

[14] Michael Handel and Lee Mosher, *Subgroup classification in* Out(F_n), preprint (2009), arXiv:0908.1255v1.

[15] W. J. Harvey, *Boundary structure of the modular group*, **97** (1981), 245–251.

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

[17] Atsushi Ishida, *The structure of subgroup of mapping class groups generated by two Dehn twists*, Proc. Japan Acad. Ser. A Math. Sci. **72** (1996), no. 10, 240–241.

[18] Nikolai V. Ivanov, *The rank of Teichmüller modular groups*, Mat. Zametki **44** (1988), no. 5, 636–644, 701.

[19] ――――, *Subgroups of Teichmüller modular groups*, Translations of Mathematical Monographs, vol. 115, American Mathematical Society, Providence, RI, 1992, Translated from the Russian by E. J. F. Primrose and revised by the author.

[20] V. A. Kaĭmanovich and A. M. Vershik, *Random walks on discrete groups: boundary and entropy*, Ann. Probab. **11** (1983), no. 3, 457–490.

[21] Richard P. Kent, IV and Christopher J. Leininger, *Shadows of mapping class groups: capturing convex cocompactness*, Geom. Funct. Anal. **18** (2008), no. 4, 1270–1325.

[22] ――――, *Uniform convergence in the mapping class group*, Ergodic Theory Dynam. Systems **28** (2008), no. 4, 1177–1195.

[23] Harry Kesten, *Symmetric random walks on groups*, Trans. Amer. Math. Soc. **92** (1959), 336–354.

[24] Daan Krammer, *Braid groups are linear*, Ann. of Math. (2) **155** (2002), no. 1, 131–156.

[25] Chris Leininger, *Graphs of Veech groups*, in preparation.

[26] Johanna Mangahas, *Uniform uniform exponential growth of subgroups of the mapping class group*, Geom. Funct. Anal. **19** (2010), no. 5, 1468–1480.

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

[28] ――――, *Geometry of the complex of curves. II. Hierarchical structure*, Geom. Funct. Anal. **10** (2000), no. 4, 902–974.

[29] John McCarthy, *A "Tits-alternative" for subgroups of surface mapping class groups*, Trans. Amer. Math. Soc. **291** (1985), no. 2, 583–612.

[30] Yair N. Minsky, *The classification of Kleinian surface groups, I: Models and bounds*, preprint (2003), arXiv:math/0302208v3.

[31] Lee Mosher, personal communication, MSRI Course on Mapping Class Groups, Fall 2007.

[32] Luis Paris and Dale Rolfsen, *Geometric subgroups of mapping class groups*, J. Reine Angew. Math. **521** (2000), 47–83.

[33] Robert C. Penner, *A construction of pseudo-Anosov homeomorphisms*, Trans. Amer. Math. Soc. **310** (1988), no. 1, 179–197.

[34] Fathi Laudenbach Poenaru, *Travaux de Thurston sur les surfaces*, Astérisque, vol. 66, Société Mathématique de France, Paris, 1979, Séminaire Orsay, With an English summary.

[35] Peter B. Shalen and Philip Wagreich, *Growth rates, Z_p-homology, and volumes of hyperbolic 3-manifolds*, Trans. Amer. Math. Soc. **331** (1992), no. 2, 895–917.

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

[37] John S. Wilson, *On exponential growth and uniformly exponential growth for groups*, Invent. Math. **155** (2004), no. 2, 287–303.

[38] Wolfgang Woess, *Random walks on infinite graphs and groups*, Cambridge Tracts in Mathematics, vol. 138, Cambridge University Press, Cambridge, 2000.