Asymptotic geometry of the mapping class group and Teichmüller space

JASON A BEHRSTOCK

In this work, we study the asymptotic geometry of the mapping class group and Teichmuller space. We introduce tools for analyzing the geometry of "projection" maps from these spaces to curve complexes of subsurfaces; from this we obtain information concerning the topology of their asymptotic cones. We deduce several applications of this analysis. One of which is that the asymptotic cone of the mapping class group of any surface is tree-graded in the sense of Druțu and Sapir; this tree-grading has several consequences including answering a question of Druțu and Sapir concerning relatively hyperbolic groups. Another application is a generalization of the result of Brock and Farb that for low complexity surfaces Teichmüller space, with the Weil–Petersson metric, is δ–hyperbolic. Although for higher complexity surfaces these spaces are not δ–hyperbolic, we establish the presence of previously unknown negative curvature phenomena in the mapping class group and Teichmüller space for arbitrary surfaces.

1 Introduction

1.1 Overview of main results

In studying the geometry and dynamics of a surface S, some of the most important mathematical objects are its mapping class group $\mathcal{MCG}(S)$, its Teichmüller space, and the complex of curves. We prove new results about the asymptotic geometry of these objects through a study of their asymptotic cones.

Recently, Druțu and Sapir showed that any group which is (strongly) relatively hyperbolic has cut-points in its asymptotic cone [19]. They asked whether this property of having asymptotic cut-points gives a characterization of such groups. We prove the following.

Theorem 7.1 *If S is any surface and $\mathrm{Cone}_\omega(\mathcal{MCG}(S))$ is any asymptotic cone, then every point of $\mathrm{Cone}_\omega(\mathcal{MCG}(S))$ is a global cut-point.*

Similarly one obtains global cut-points in each asymptotic cone of Teichmüller space with the Weil–Petersson metric.

This result answers the question of [19] in the negative for, as shown by Anderson, Aramayona and Shackleton [1] and independently by the author with Druţu and Mosher [4], the group $\mathcal{MCG}(S)$ is not hyperbolic relative to any family of finitely generated subgroups. This establishes the mapping class groups as the first known examples of groups which are not relatively hyperbolic, but whose asymptotic cones are tree-graded. See [4] for other applications of Theorem 7.1 and its Corollary 7.3.

We obtain the following more specific result on the asymptotic geometry of the mapping class group.

Theorem 6.5 *In any surface, S, and any asymptotic cone, $\mathrm{Cone}_\omega(\mathcal{MCG}(S))$, there exists an isometrically embedded $\mathbb{R}$–tree, $F_0 \subset \mathrm{Cone}_\omega(\mathcal{MCG}(S))$ with the following properties:*

- *F_0 contains infinitely many distinct infinite geodesics arising as asymptotic cones of axes of pseudo-Anosovs;*

- *for each pair of points on F_0, there exists a unique topological path in $\mathrm{Cone}_\omega(\mathcal{MCG}(S))$ connecting them and this path is contained in F_0.*

A similar result holds replacing the mapping class group by Teichmüller space with the Weil–Petersson metric.

One application of Theorem 6.5 is that it shows that the directions of pseudo-Anosov axes are negatively curved, ie, in the direction of each pseudo-Anosov the divergence function is superlinear. This readily implies that these directions are quasi-geodesically stable, ie, for the quasi-geodesic axis given by taking powers of a pseudo-Anosov, there is a bound $M(K, C)$ for which any (K, C)–quasigeodesic with endpoints on this path must be completely contained within the M neighborhood of the axis. We also establish that the Weil–Petersson metric on Teichmüller space is negatively curved in the directions of geodesics all of whose subsurface projections are uniformly bounded. Thus although for higher complexity surfaces the Weil–Petersson metric is not δ–hyperbolic, this result shows that there are directions in which one gets the useful consequences of δ–hyperbolicity. (Note that such a result does not follow from the negative sectional curvature of the Weil–Petersson metric since these curvatures are not bounded away from 0.) Results on stability of certain families of geodesics in the Teichmüller metric was proved by Minsky [45], thus our results complete this analogy between these three spaces. (After completing this result, we were informed by Minsky

that in current joint work with Brock and Masur they will give an alternate proof of the stability of a similar family of Weil–Petersson geodesics.)

When the complexity of the surface is small enough, we recover the following result which was first established by Brock and Farb [16]. (After this work was completed another alternate proof of this result appeared using interesting techniques involving the geometry of $CAT(0)$ spaces, see Aramayona [2].)

Theorem 5.1   *With the Weil–Petersson metric, the Teichmüller spaces for the surfaces $S_{1,2}$ and $S_{0,5}$ are each δ–hyperbolic.*

This result implies that any subset of an asymptotic cone of Teichmüller space of $S_{1,2}$ or $S_{0,5}$ contains a point which disconnects it. By contrast, the asymptotic cone of the Teichmüller space of any surface of higher complexity contains bi-Lipschitz embeddings of $\mathbb{R}^n$ with $n \geq 2$, these subsets can not be disconnected by any point. Nonetheless, Theorem 7.1 implies that each point is a global cut-point; this is one of several ways to consider Theorem 7.1 as a generalization of Theorem 5.1.

The original motivation for this work was to study quasiflats in the mapping class group. In particular, it had been conjectured by Brock–Farb that the mapping class group admits quasiflats only up to the maximal dimension for which it contains abelian subgroups (by work of Birman, Lubotzky and McCarthy [8] the latter number is known to be $3g + p - 3$). Using Theorem 6.5 as the base case for an inductive argument, the author in joint work with Minsky [6] have proven this "Rank Conjecture" (using different techniques Hamenstädt [25] has also obtained a proof of this result).

1.2   Methods

The above results are produced by analyzing the *complex of curves*, a simplicial complex encoding intersection patterns of curves on a surface. This complex was first introduced by Harvey and has since been used to prove many deep results about the mapping class group (see Harer [26], Harvey [27], Hatcher [28], and Ivanov [33]). More recently Masur and Minsky proved δ–hyperbolicity of the complex of curves [42], and then used this to establish new methods for studying the mapping class group [41]. The complex of curves has also found many applications to the study of 3–manifolds; in particular it has proved crucial in the recent proof of the Ending Lamination Conjecture (see Brock, Canary and Minsky [44; 15]). For other applications see: Behrstock and Margalit [5], Bestvina and Fujiwara [7], Bowditch [11], Hempel [29], Irmak [32] and Margalit [40].

For each subsurface $X \subset S$, we consider a "projection" map

$$\pi_X \colon \mathcal{MCG}(S) \to \mathcal{C}(X)$$

from the mapping class group of S into the curve complex of X, which is closely related to the map taking a curve in S to its intersection with X. The vertices of the curve complex of a surface, $\mathcal{C}(X)$, are homotopy classes of essential, non-peripheral closed curves on X. The distance between two distinct homotopy classes of curves, $\mu, \nu \in \mathcal{C}(X)$, is a measure of how complicated the intersection of representatives of these curves must be. For example as long as X has genus larger than one: $d_X(\mu, \nu) = 1$ if and only if μ and ν can be realized disjointly on X whereas $d_X(\mu, \nu) \geq 3$ if and only if μ and ν fill X (*fill* means that for any representatives of μ and ν the set $X \setminus (\mu \cup \nu)$ consists of disks and once-punctured disks).

The following tool provides the starting point for our analysis.

Theorem 4.3 (Projection estimates) *Let Y and Z be two overlapping subsurfaces of S. Then for any $\mu \in \mathcal{MCG}(S)$:*

$$d_{\mathcal{C}(Y)}(\partial Z, \mu) > M \implies d_{\mathcal{C}(Z)}(\partial Y, \mu) \leq M.$$

Where M depends only on the topological type of the surface S.

Combined with work of Masur–Minsky [41] this result is useful in studying the metric geometry of mapping class groups and Teichmüller space via their subsurface projections. (Theorem 4.3 can be generalized to study projections of the mapping class group into a product of curve complexes. Doing so yields an interesting picture of certain aspects of the asymptotic geometry of the mapping class group, this direction is developed further by the author in [3].) Our study of the metric geometry of mapping class groups and Teichmüller space is done via an analysis of their asymptotic cones. We define a projection map to send elements of the mapping class group of a surface S to the mapping class group of a subsurface $Z \subset S$. This projection map has several nice properties, for instance it sends points to uniformly bounded diameter subsets (Lemma 6.1), and is used in defining an important new object introduced in this paper: the *strongly bounded subset*, F_0, of the asymptotic cone. This subset is part of the *transversal tree* of $\mathrm{Cone}_\omega \mathcal{MCG}(S)$ and plays a large role in the Behrstock–Minsky proof of the Rank Conjecture [6].

1.3 Outline of subsequent sections

In Section 2 we discuss the tools of asymptotic geometry which are fundamental to the constructions and results in the rest of this work. All the results in this section are well

known and thus we are as much setting up notation as we are reminding the reader of which tools to keep at their fingertips for the remainder of this work.

Surfaces are the main characters throughout this work. In Section 3 we discuss the marking complex, which is our quasi-isometric model of choice for dealing with the mapping class group. We also remind the reader of various facts we will need concerning the mapping class group, the complex of curves, and Teichmüller space.

The original material in this work begins with the results in Section 4. Here we calculate estimates on the image of the "projection" map from the marking complex into the product of the curve complexes of the constituent subsurfaces. In this section we produce our starting point, Theorem 4.3 (Projection estimates), which is a technical result we use in our analysis of the mapping class group. In Theorem 4.4 (Projection estimates; geometric version) we provide a geometric interpretation of Theorem 4.3.

The final three sections apply the Projection estimates Theorem and related techniques to various situations. In the first of these we (re)prove hyperbolicity of the Weil–Petersson metric on Teichmüller space and of the mapping class group in some low complexity cases, Theorems 5.1 and 5.2. The techniques introduced in this section provide a good warm-up for those used to prove the key step in its generalization, Theorem 6.5.

In Section 6 we generalize the hyperbolicity of low complexity mapping class groups and Teichmüller spaces to deduce the existence of negatively curved directions in the cases of higher complexity surfaces. The main result of this section is Theorem 6.5 which identifies some new aspects of the geometry of $\mathrm{Cone}_\omega(\mathcal{MCG})$ which we exploit in the following section.

We end with Section 7 where we relate our earlier work to questions about relative hyperbolicity in the mapping class group. Here we answer some questions about mapping class groups and tree-graded spaces and raise some new question.

Acknowledgments

2 Background I: Large scale geometry

2.1 Geometric Group Theory

There are several equivalent definitions for hyperbolicity; throughout this paper we will have in mind Rips' *thin triangle* characterization:

Definition 2.1 A geodesic metric space X is called δ–*hyperbolic* if there exists a constant δ so that for each triple $a, b, c \in X$ and each choice of geodesics $[a, b], [b, c]$, and $[a, c]$ one has that $[a, b]$ is contained in a δ–neighborhood of the union of $[a, c]$ and $[b, c]$.

Note that we write $[a, b]$ to denote a geodesic between a and b even though this choice need not be unique.

Definition 2.2 Let (X, d_X) and (Y, d_Y) be metric spaces. A map $\phi\colon X \to Y$ is called a (K, C)–*quasi-isometric embedding* if there exist constants $K \geq 1$ and $C \geq 0$ such that for all $a, b \in X$

$$\frac{1}{K} d_X(a, b) - C \leq d_Y(\phi(a), \phi(b)) \leq K d_X(a, b) + C.$$

It is a (K, C)–*quasi-isometry* if additionally, ϕ has the property that every point of Y lies in the C–neighborhood of $\phi(X)$.

When there exists a (K, C)–quasi-isometry $\phi\colon X \to Y$ between two metric spaces X and Y, we say that they are (K, C)–*quasi-isometric* and denote this $X \sim_{K,C} Y$; When the choice of constants is not important (as will often be the case), we simply say X and Y are *quasi-isometric*, denoted $X \sim Y$.

For geodesic spaces hyperbolicity is a quasi-isometry invariant; note that the hyperbolicity constant may change (see Ghys and de la Harpe [21]).

In order to apply this definition to groups, we recall the natural left invariant metric on a group G with finite generating set S. The *word metric on G relative to the generating set S* is given by $d_{G,S}(a, b) = |a^{-1}b|$ where $|a^{-1}b|$ is the smallest number of letters needed to represent the group element $a^{-1}b$ in terms of letters from S and their inverses. Although this metric depends on the choice of generating set, for finitely generated groups, its quasi-isometry type does not. Thus since hyperbolicity is a quasi-isometry invariant, the notion of a group being *hyperbolic* is well defined.

A basic non-hyperbolic example where we will use the notion of quasi-isometries is $\mathbb{Z}^n$ (with $n > 1$), which we (easily) observe is quasi-isometric to $\mathbb{R}^n$. In any metric

space X one can study quasi-isometric embeddings of $\mathbb{R}^n$ into X, these are called *quasiflats of rank n*. Generalizing the above Euclidean example, one way to quantify non-hyperbolicity of a metric space X is by calculating its *geometric rank*, namely the largest rank of any quasiflat in X. Having geometric rank greater than one is an obstruction to hyperbolicity, as it forces X to have arbitrarily fat triangles.

We now record a technical lemma which will be useful in Section 5. The proof is an exercise in manipulating thin triangles (see [3] for a proof). When the ambient space is clear we use the notation $B_\delta(x)$ to denote the closed ball of radius δ around a point x.

Lemma 2.3 Let Z be a δ–*hyperbolic space and* γ *a geodesic in* Z. *If* $\phi\colon Z \to 2^\gamma$ *sends each point of* Z *to the set of closest points on* γ, *then for each point* $z \in Z$ *we have* $\mathrm{diam}(\phi(z)) \le 4\delta + 2$.

2.2 Asymptotic Cones

Ultrafilters and ultraproducts have been used in logic since Gödel introduced them in the 1930's. Among other applications they can be used to give a slick proof of the compactness theorem of first order logic and to rigorize calculus, *a la* Robinson's non-standard analysis. Van den Dries and Wilkie introduced these tools to topologists when they reformulated (and slightly strengthened) Gromov's famous Polynomial growth Theorem using this language (compare Gromov [24] and van den Dries and Wilkie [53]). These techniques streamlined parts of the proof, replacing an iterated procedure of passing to subsequences with one ultralimit, and broadened the applicability of Gromov's construction, Gromov's original definition of asymptotic cone only applied to nilpotent groups whereas the ultrafied definition works for any finitely generated group. (For other applications of asymptotic cones in geometric group theory see: Burillo [17], Bridson [12], Druțu[18], Druțu and Sapir [19], Kapovich and Leeb [36] Kleiner and Leeb [38], Kramer, Shelah, Tent and Thomas [39] and Riley [50].)

A *non-principal ultrafilter* on the integers, denoted ω, is a nonempty collection of sets of integers with the following properties:

(1) If $S_1 \in \omega$ and $S_2 \in \omega$, then $S_1 \cap S_2 \in \omega$.

(2) If $S_1 \subset S_2$ and $S_1 \in \omega$, then $S_2 \in \omega$.

(3) It is *maximal* in the sense that for each $S \subset \mathbb{Z}$ exactly one of the following must occur: $S \in \omega$ or $\mathbb{Z} \setminus S \in \omega$.

(4) ω does not contain any finite sets. (This is the *non-principal* aspect.)

Since we are only concerned with non-principal ultrafilter, we will simply refer to these as *ultrafilters*.

An equivalent way to view an ultrafilter is as a finitely additive probability measure, with no atoms, defined on all subsets of integers and which takes values in $\{0, 1\}$. This way of thinking is consistent with the intuition that the ultrafilter sees the behavior on only the large subsets of $\mathbb{N}$ and when this is done one says that ω–almost every integer has a given property when the set of integers with this property is in ω.

For an ultrafilter ω, a metric space (X, d), and a sequence of points $\langle y_i \rangle_{i \in \mathbb{N}}$ we define y to be the *ultralimit of* $\langle y_i \rangle_{i \in \mathbb{N}}$ *with respect to* ω, denoted $y = \lim_\omega y_i$, if and only if for all $\epsilon > 0$ one has $\{i \in \mathbb{N} : d(y_i, y) < \epsilon\} \in \omega$. When the ultralimit of a sequence exists it is a unique point; further, it is easy to see that when the sequence $\langle y_i \rangle_{i \in \mathbb{N}}$ converges, then the ultralimit is the same as the ordinary limit. It is also worth noting that in a compact metric space ultralimits always exist.

Fix an ultrafilter ω and a family of based metric spaces (X_i, x_i, d_i). Using the ultrafilter, a pseudo-distance on $\prod_{i \in \mathbb{N}} X_i$ is provided by:

$$d_\omega(\langle a_i \rangle, \langle b_i \rangle) = \lim_\omega d_{X_i}(a_i, b_i) \in [0, \infty].$$

The *ultralimit of* (X_i, x_i) is then defined to be:

$$\lim_\omega (X_i, x_i) = \{y \in \prod_{i \in \mathbb{N}} X_i : d_\omega(y, \langle x_i \rangle) < \infty\}/ \sim,$$

where for two points $y, z \in \prod_{i \in \mathbb{N}} X_i$ we define $y \sim z$ if and only if $d_\omega(y, z) = 0$.

The pseudo-metric d_ω takes values in $[0, \infty]$—considering only those points in $\lim_\omega(X_i, x_i)$ which are a finite distance from the basepoint restricts our attention to a maximal subset where the pseudo-distance does not obtain the value ∞. The relation $\sim$ quotients out points whose pseudo-distance from each other is zero; in an analogy to measure theory we think of this process as identifying sequences which agree almost everywhere. These two conditions combine to make $d_\omega(y, z)$ into a metric.

Just as ultralimits for sequences of points generalize the topological notion of limit, ultralimits of sequences of metric spaces generalize Hausdorff limits of metric spaces. In particular, for a Hausdorff precompact family of metric spaces the ultralimit of the sequence is a limit point with respect to the Hausdorff topology (see [38] for a proof of this and other related facts).

Definition 2.4 The *asymptotic cone of $(X, \langle x_i \rangle, \langle d_i \rangle)$ relative to the ultrafilter ω* is defined by:

$$\mathrm{Cone}_\omega(X, \langle x_i \rangle, \langle d_i \rangle) = \lim_\omega (X, x_i, \frac{1}{d_i}).$$

When it is not a source of confusion we tend to suppress writing the basepoint. Note that for for homogeneous spaces (like Cayley graphs), the asymptotic cone does not depend on the basepoint $\langle x_i \rangle$ and thus we shall always take this to be $\langle 1 \rangle$ and drop it from the notation.

When taking such ultralimits, we often use the notation $\lim_\omega \frac{1}{d_i}(x_\alpha, 0_\alpha)$, when there is an implied ambient space T_α with basepoint 0_α, in order to emphasize that we are looking at an ultralimit in $\mathrm{Cone}_\omega(T_\alpha, 0_\alpha)$.

When considering the asymptotic cone, especially for the first time, the following quotation from Thomas Hobbes (1588–1679), will likely seem relevant.

> To understand this for sense it is not required that a man should be a geometrician or a logician, but that he should be mad.

Asymptotic cones provide a way to replace a metric space with a "limiting" space which carries information about sequences in the original space which leave every compact set. This process encodes the asymptotic geometry of a space into standard algebraic topology invariants of its asymptotic cone.

A common use of asymptotic cones is in their relation to hyperbolicity as demonstrated by the next result and Theorem 2.7. First we give a preliminary definition.

Definition 2.5 An $\mathbb{R}$–tree is a metric space (X, d) such that between any two points $a, b \in X$ there exists a unique topological arc γ connecting them and γ is isometric to the interval $[0, d_X(a, b)] \subset \mathbb{R}$.

Proposition 2.6 *For a sequence $\delta_i \to 0$, the ultralimit of δ_i–hyperbolic spaces is an $\mathbb{R}$–tree.*

In particular, if X is a δ–hyperbolic space, then $\mathrm{Cone}_\omega X$ is an $\mathbb{R}$–tree.

If X is a δ–hyperbolic space which is sufficiently complicated (eg, the cayley graph of a non-elementary hyperbolic group) then $\mathrm{Cone}_\omega X$ is an $\mathbb{R}$–trees with uncountable branching at every point.

A partial converse to Proposition 2.6 is provided by the following well known result. The statement of this theorem first appeared, without proof, in Gromov [22]. A proof can be found in Druţu [18].

Theorem 2.7 *If every asymptotic cone of a metric space X is an $\mathbb{R}$–tree, then X is δ–hyperbolic.*

The following is an easy observation that simplifies the situation when dealing with finitely many equivalence classes of Seq, where Seq is the set of sequences in a countable alphabet considered up to the equivalence relation they they agree ω–almost everywhere. When we consider Seq in the sequel the alphabet will consist of subsurfaces of a given surface.

Lemma 2.8 *For any finite set $\Gamma \subset$ Seq, the elements of Γ are pairwise distinct in Seq if and only if there exists a set $K \in \omega$ for which each $\gamma, \gamma' \in \Gamma$ has $\gamma_i \neq \gamma'_i$ for every $i \in K$*

Proof Fix a finite set $\Gamma \subset$ Seq.

($\Longrightarrow$) The maximal clause in the definition of ultrafilter states that for each $K \subset \mathbb{N}$ either $K \in \omega$ or $\mathbb{N} \setminus K \in \omega$. We suppose that for each pair of elements $\gamma, \gamma' \in \Gamma$ we have $\gamma \sim \gamma'$ and thus the set K of indices where $\gamma_i = \gamma'_i$ must have $K \notin \omega$. Maximality of the ultrafilter then implies that $\mathbb{N} \setminus K \in \omega$. Thus for each $i \in \mathbb{N} \setminus K \in \omega$ we have $\gamma_i \neq \gamma'_i$, we define $\mathbb{N} \setminus K = K'_{\gamma, \gamma'}$.

Since ultrafilters are closed under finite intersections, the intersection over all pairs γ, γ' of $K_{\gamma, \gamma'}$ yields a set $J \in \omega$ where $\gamma_i \neq \gamma'_i$ for every $i \in J$ and every $\gamma, \gamma' \in \Gamma$.

($\Longleftarrow$) Let $K \in \omega$ be the set of indices for which $\gamma_i \neq \gamma'_i$ for each $i \in K$ and $\gamma, \gamma' \in \Gamma$. Then for any γ, γ' we have $\{i : \gamma_i = \gamma'_i\} \in \mathbb{N} \setminus K$ and hence is not in ω. Thus $\gamma \sim \gamma'$. $\square$

We end this section with a summary of some standard results about asymptotic cones that we will use in the sequel (see Bridson and Haefliger [13], Kleiner and Leeb [38] or Kapovich [34]).

Proposition 2.9 *Fix a non-principal ultrafilter ω.*

(1) $\text{Cone}_\omega(X)$ *is a complete metric space.*

(2) $\text{Cone}_\omega(X_1 \times X_2) = \text{Cone}_\omega(X_1) \times \text{Cone}_\omega(X_2)$

(3) $\text{Cone}_\omega \mathbb{R}^n = \mathbb{R}^n$

(4) *The asymptotic cone of a geodesic space is a geodesic space.*

(5) *A (K, C)–quasi-isometry between metric spaces induces a K–bi-Lipschitz map between their asymptotic cones.*

3 Background II: Surfaces and related structures

3.1 Complex of Curves

We use S to denote a connected, orientable surface of genus $g = g(S)$ with $p = p(S)$ punctures. (Note that we will interchangeably use the words puncture and boundary, as the distinction does not affect the results contained in this work.) We use the terms *subsurface* and *domain* to refer to a homotopy class of an essential, non-peripheral, connected subsurface of S (subsurfaces are not assumed to be proper unless explicitly stated). When we refer to the boundary of a domain this will mean the collection of (homotopy classes of) non-peripheral closed curves which bound the domain as a subset of S.

We now recall the construction of the complex of curves and relevant machinery developed by Masur and Minsky which we will use in our study; for further details consult [41]. We use $\xi(S) = 3g(S) + p(S) - 3$ to quantify the complexity[1] of the surface S. Recall that when positive, $3g(S) + p(S) - 3$ is the maximal number of disjoint homotopy classes of essential and non-peripheral simple closed curves which can be simultaneously realized on S. Naturality of $\xi(S)$ as a measure of complexity is justified by the property that it decreases when one passes from a surface to a proper subsurface (recall our convention of considering surfaces up to homotopy). Since our interest is in hyperbolic surfaces and their subsurfaces, we only consider subsurfaces with $\xi > -2$ (thus ignoring the disk and the sphere). Additionally, as it is not a hyperbolic surface nor does it appear as a subsurface of any hyperbolic surface we will usually ignore $S_{1,0}$ (although much of our discussion has analogues for this case); thus $\xi = 0$ is used only to denote the thrice punctured sphere.

Introduced by Harvey [27] to study the boundary of Teichmüller space, the complex of curves has proven to be a useful tool in the study of Teichmüller space, mapping class groups, and 3–manifolds. The complex of curves is a finite dimensional complex which encodes information about the surface via the combinatorics of simple closed curves. Analysis using the curve complex is necessarily delicate since the complex is locally infinite except in a few low genus cases.

Definition 3.1 The *complex of curves for* S, denoted $\mathcal{C}(S)$, consists of a *vertex* for every homotopy class of a simple closed curve which is both non-trivial and non-peripheral. The N–*simplices* of $\mathcal{C}(S)$ are given by collections of $N + 1$ vertices whose homotopy classes can all simultaneously be realized disjointly on S.

---

[1]In [41] they use $\xi(S) = 3g(S) + p(S)$, but we use the current version because it has better additive properties when generalized to disconnected surfaces, as we need in Section 6.

This definition works well when $\xi(S) > 1$, however it must be modified slightly for the surfaces of small complexity, which we refer to as the *sporadic cases*. The two cases where $\xi(S) = 1$ are $S_{1,1}$ and $S_{0,4}$ (the following discussion works for $S_{1,0}$ as well); here any two distinct homotopy classes of essential simple closed curves must intersect (at least once in $S_{1,1}$ and twice in $S_{0,4}$). In these cases the above definition of $\mathcal{C}(S)$ would have no edges, only vertices. Accordingly, we modify the definition so that an edge is added between two homotopy classes when they intersect the minimal amount possible on the surface (ie, once for $S_{1,1}$ and twice for $S_{0,4}$). With this definition $\mathcal{C}(S_{1,1})$ and $\mathcal{C}(S_{0,4})$ are each connected, indeed they are each isometric to the classical Farey graph.

When $\xi(S) = 0, -2$, or -3 then $\mathcal{C}(S)$ is empty. The final modification we make is for the case where we have $A \subset S$ with $\xi(A) = -1$, the annulus. The annulus doesn't support a finite area hyperbolic metric, so our interest in it derives from the fact that it arises as a subsurface of hyperbolic surfaces. Indeed, annuli will play a crucial role as they will be used to capture information about Dehn twists. Given an annulus $A \subset S$, we define $\mathcal{C}(A)$ to be based homotopy classes of arcs connecting one boundary component of the annulus to the other. More precisely, denoting by $\tilde{A}$ the annular cover of S to which A lifts homeomorphically, we use the compactification of $\mathbb{H}^2$ as the closed unit disk to obtain a closed annulus $\hat{A}$. We define the vertices of $\mathcal{C}(A)$ to be homotopy classes of paths connecting one boundary component of $\hat{A}$ to the other, where the homotopies are required to fix the endpoints. Edges of $\mathcal{C}(A)$ are pairs of vertices which have representatives with disjoint interiors. Giving edges a Euclidean metric of length one, in [41] it is proven that $\mathcal{C}(A)$ is quasi-isometric to $\mathbb{Z}$.

The following foundational result, which is the main theorem of [42], will be useful in our later analysis. For another proof see Bowditch [9], which provides a constructive computation of a bound on the hyperbolicity constant.

Theorem 3.2 (Hyperbolicity of $\mathcal{C}(S)$ [42; 9]) *For any surface S, $\mathcal{C}(S)$ is an infinite diameter δ–hyperbolic space (as long as it is non-empty).*

Throughout this paper we use the convention that "intersection" refers to transverse intersection. Thus for example if we consider a subsurface $Y \subsetneq S$ and an element $\gamma \in \partial Y$ which is not homotopic to a puncture of S then we consider the annulus around γ to not intersect Y.

For a surface with punctures one can consider the *arc complex* $\mathcal{C}'(S)$, which is a close relative of the curve complex. When $\xi(S) > -1$ we define the vertices[2] of $\mathcal{C}'(S)$

---

[2]This definition differs from that given by Harer [26] where an arc complex is considered consisting of only arcs and thus does not contain $\mathcal{C}(S)$ as a subcomplex. Our definition agrees with that in [41].

(denoted $\mathcal{C}'_0(S)$) to consist of elements of $\mathcal{C}_0(S)$ as well as homotopy classes of simple arcs on S with endpoints lying on punctures of S, which don't bound a disk or a once punctured disk on either side. As done for the curve complex, we define N–simplices of $\mathcal{C}'(S)$ to be collections of $N + 1$ vertices which can simultaneously be realized on the surface as disjoint arcs and curves. For annuli we define $\mathcal{C}'(A) = \mathcal{C}(A)$.

The arc complex arises naturally when one tries to "project" an element $\gamma \in \mathcal{C}(S)$ into $\mathcal{C}(Y)$ where $Y \subset S$. When $\xi(Y) > 0$ we define

$$\pi'_Y \colon \mathcal{C}_0(S) \to 2^{C'_0(Y)}$$

by the following:

- If $\gamma \cap Y = \varnothing$, then define $\pi'_Y(\gamma) = \varnothing$.
- If $\gamma \subset Y$, then define $\pi'_Y(\gamma) = \{\gamma\}$.
- If $\gamma \cap \partial Y \neq \varnothing$, then after putting γ in a position so it has minimal intersection with ∂Y we identify parallel arcs of $\gamma \cap Y$ and define $\pi'_Y(\gamma)$ to be the union of these arcs and any closed curves in $\gamma \cap \partial Y \neq \varnothing$.

In the last case, it follows from the definition of $\mathcal{C}'(Y)$ that $\pi'_Y(\gamma)$ is a subset of $\mathcal{C}'_0(Y)$ with diameter at most one. Thus whenever $\pi'_Y(\gamma) \neq \varnothing$, it is a subset of $\mathcal{C}'(Y)$ of diameter at most one. Moreover, when $\xi(S) > 0$, there is a map $\phi_Y \colon \mathcal{C}'_0(Y) \to 2^{\mathcal{C}_0(Y)}$, which implies that the arc complex embeds into the curve complex as a cobounded subset—the map ϕ_Y sends each arc to the boundary curves of a regular neighborhood of its union with ∂Y. For $Y \subset S$ with $\xi(Y) > 0$ we define $\pi_Y = \phi_Y \circ \pi'_Y \colon \mathcal{C}(S) \to \mathcal{C}(Y)$ (this map is actually to the power set of $\mathcal{C}(Y)$, but for the sake of readability and since the map is to bounded sets, we abuse notation and simply write $\mathcal{C}(Y)$).

When $\xi(Y) = -1$ any curve γ which crosses Y transversally has a lift $\widetilde{\gamma} \in \widetilde{Y}$ with at least one component which connects the two boundary components of $\widehat{Y}$. Together, the collection of lifts which connect the boundary components form a finite subset of $\mathcal{C}'(Y)$ with diameter at most 1, define $\pi_Y(\gamma)$ to be this set. If γ doesn't intersect Y or is the core curve of Y, then define $\pi_Y(\gamma) = \varnothing$. Also, for consistency we define $\pi_Y \colon v \to \{v\}$ for $v \in \mathcal{C}_0(Y)$.

Since we often work with subsets of $\mathcal{C}(Y)$ rather than points, we use the following notation. For a set valued map $f \colon X \to 2^Y$ and a set $A \subset X$ we define $f(A) = \cup_{x \in A} f(x)$. If $X \subseteq S$ and $\mu, \nu \in \mathcal{C}(S)$ we adopt the notation

$$d_X(\mu, \nu) = d_{\mathcal{C}(X)}(\pi_X(\mu), \pi_X(\gamma)).$$

Further, given sets $A, B \in \mathcal{C}(S)$ we set $d_X(A, B) = \min\{d_X(\alpha, \beta) : \alpha \in A \text{ and } \beta \in B\}$. Also, we write $\operatorname{diam}_X(A)$ to refer to the diameter of the set $\pi_{\mathcal{C}(X)}(A)$ and $\operatorname{diam}_X(A, B)$

for $\text{diam}_{\mathcal{C}(X)}(A \cup B)$ in order to emphasize the symmetry between our use of minimal distance and diameter. (These same conventions apply when considering markings, as defined in the next section.)

An extremely useful result concerning these projections is the following (see [41] for the original proof, and [44] where the bound is corrected from 2 to 3):

Lemma 3.3 (Lipschitz projection [41]) *Let Y be a subdomain of S. For any simplex ρ in $\mathcal{C}(S)$, if $\pi_Y(\rho) \neq \varnothing$ then $\text{diam}_Y(\rho) \leq 3$.*

In light of this lemma, for any pair of domains $Y \subset Z$ we consider $\mathcal{C}(Z) \setminus B_1(\partial Y)$ with the path metric, ie, the distance between $\gamma, \gamma' \in \mathcal{C}(Z) \setminus B_1(\partial Y)$ is the length of the shortest path connecting them in $\mathcal{C}(Z) \setminus B_1(\partial Y)$. With this metric there is a *coarsely Lipschitz* map $\pi_{Z \to Y} : \mathcal{C}(Z) \setminus B_1(\partial Y) \to \mathcal{C}(Y)$, where a map is defined to be $(K, C)-$ *coarsely Lipschitz* if and only if there exist a pair of constants K, C such that for each $a, b \in \mathcal{C}(Z) \setminus B_1(\partial Y)$ we have $d_Y(\pi_{Z \to Y}(a), \pi_{Z \to Y}(b)) \leq K d_{\mathcal{C}(Z) \setminus B_1(\partial Y)}(a, b) + C$. By the above lemma and the fact that $\pi_{Z \to Y}$ sends points of $\mathcal{C}(Z) \setminus B_1(\partial Y)$ to subsets of $\mathcal{C}(Y)$ of diameter at most 3, we have the following which we shall use in Section 4. (See Figure 1 for a cartoon of the behavior of this map near ∂Y.)

Corollary 3.4 *For any domain $Y \subset Z$, endowing $\mathcal{C}(Z) \setminus B_1(\partial Y)$ with the path metric, we have*

$$\pi_{Z \to Y} : \mathcal{C}_1(Z) \setminus B_1(\partial Y) \to \mathcal{C}(Y)$$

is coarsely Lipschitz (with constants $K = 3$ and $C = 3$).

In Figure 1, we observe that although $\pi_{Z \to Y}$ is Lipschitz, pairs of points like γ_1 and γ_5, although distance 2 in $\mathcal{C}(Z)$, can be found which are arbitrarily far apart in the path metric on $\mathcal{C}(Z) \setminus B_1(\partial Y)$ and thus their distance when projected to $\mathcal{C}(Y)$ are also made arbitrarily large.

Letting $\mathcal{C}_1(S)$ denote the one-skeleton of $\mathcal{C}(S)$, we state the following theorem which is one of the key technical results from [41].

Theorem 3.5 (Bounded geodesic image [41]) *Let $Y \subsetneq S$ with $\xi(Y) \neq 0$ and let g be a geodesic segment, ray, or line in $\mathcal{C}_1(S)$, such that $\pi_Y(v) \neq \varnothing$ for every vertex v of g. There is a constant D depending only on $\xi(S)$ so that $\text{diam}_Y(g) \leq D$.*
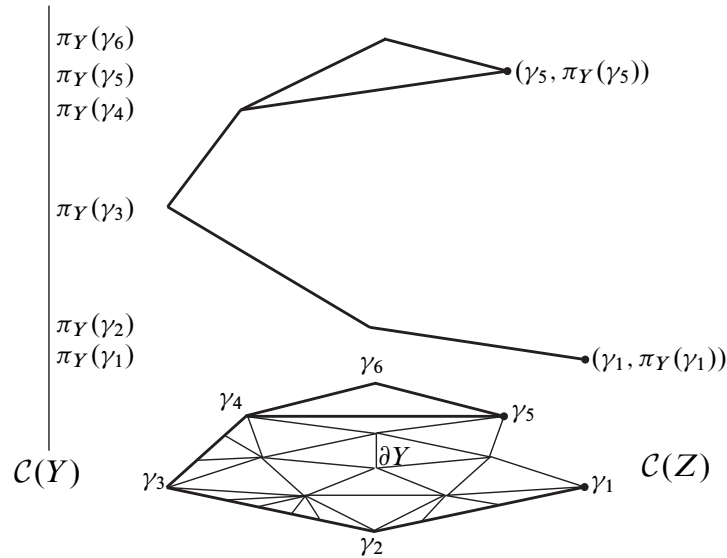
Figure 1: Letting $Y \subset Z$, the above is a caricature of the graph of $\pi_{Z \to Y} \colon \mathcal{C}(Z) \setminus B_1(\partial Y) \to \mathcal{C}(Y)$ near the point $\partial Y \in \mathcal{C}(Z)$.

3.2 Markings

In this section we describe the quasi-isometric model we will use for the mapping class group and explain the tools developed in [41] for computing with this model.

Definition 3.6 A *marking*, μ, on S is a collection of *base curves* to each of which we (may) associate a *transverse curve*. These collections are made subject to the constraints:

- The *base curves*, $\text{base}(\mu) = \{\gamma_1, \ldots, \gamma_n\}$, form a simplex in $\mathcal{C}(S)$.
- The *transverse curve*, t, associated to a given base curve γ is either empty or an element of $\mathcal{C}_0(S)$ which intersects t once (or twice if $S = S_{0,4}$) and projects to a subset of diameter at most 1 in the annular complex $\mathcal{C}(\gamma)$.

When the transverse curve t is empty we say that γ doesn't have a transverse curve.

If the simplex formed by $\text{base}(\mu)$ in $\mathcal{C}(S)$ is top dimensional and every curve has a transverse curve, then we say the marking is *complete*.

If a marking has the following two properties, then we say the marking is *clean*. First, for each $\gamma \in \text{base } \mu$ its transversal t is disjoint from the rest of the base curves. Second, for each γ and t as above, $t \cup \gamma$ fills a surface denoted $F(t, \gamma)$ with $\xi(F(t, \gamma)) = 1$ and in which $d_{\mathcal{C}(F(t,\gamma))}(t, \gamma) = 1$. (See Figure 2.)
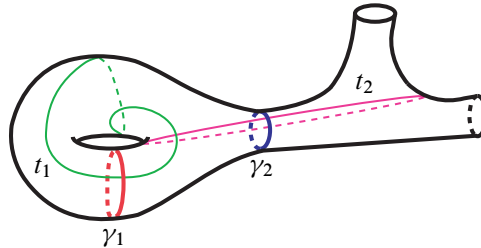
Figure 2: A complete clean marking $\mu \in \mathcal{M}(S_{1,2})$ with $\text{base}(\mu) = \{\gamma_1, \gamma_2\}$

Let μ denote a complete clean marking with pairs $(\gamma_i, t(\gamma_i))$, we take as *elementary moves* the following two relations on the set of complete clean markings:

(1) *Twist*: For some i, we replace $(\gamma_i, t(\gamma_i))$ by $(\gamma_i, t'(\gamma_i))$ where $t'(\gamma_i)$ is the result of one full (or half when possible) twist of $t(\gamma_i)$ around γ_i. The rest of the pairs are left unchanged.

(2) *Flip*: For some i we swap the roles of the base and transverse curves between γ_i and $t(\gamma_i)$. After doing this the complete marking may no longer be clean, one then needs to replace the new marking with a compatible clean one.

We say the clean marking μ is *compatible* with μ' if they have the same base curves, each base curve γ has a transverse curve t in one marking if and only if it has a transverse curve t' in the other marking, and when t exists then $d_\gamma(t, t')$ is minimal among all choices of t'.

In [41] it is shown that there exists a bound (depending only on the topological type of S) on the number of clean markings which are compatible with any other given marking. Thus even though the flip move is defined by choosing an arbitrary compatible complete clean marking, it is canonical up to some uniformly bounded amount of ambiguity.

One then defines the *marking complex*, denoted $\mathcal{M}(S)$, to be the graph formed by taking the complete clean markings on S as vertices and connecting two vertices by an edge if they differ by an elementary move. It is not hard to check that $\mathcal{M}(S)$ is a locally finite graph and that the mapping class group acts on it cocompactly and properly discontinuously. A consequence of this which we use throughout this work is:

Corollary 3.7 [41] $\mathcal{M}(S)$ *is quasi-isometric to the mapping class group.*

For any subsurface $Y \subset S$ with $\mathcal{C}(Y) \neq \varnothing$ we considered in the previous section the *subsurface projection*, $\pi_Y \colon \mathcal{C}(S) \to 2^{\mathcal{C}(Y)}$. More generally one can consider subsurface projections from the marking complex; since this definition generalizes the above map

we also denote it $\pi_Y \colon \mathcal{M}(S) \to 2^{\mathcal{C}(Y)}$. When $\mu \in \mathcal{M}(S)$ we define $\pi_Y(\mu)$ to be $\pi_Y(\mathrm{base}(\mu))$ unless Y is an annulus around a curve $\gamma \in \mathrm{base}(\mu)$, in which case we define $\pi_Y(\mu) = \pi_Y(t)$ where t is the transverse curve to γ. An important observation is that when μ is a complete marking on S then $\pi_Y(\mu) \neq \varnothing$ for each $Y \subset S$.

There is also a Lipschitz projection property for projections of markings.

Lemma 3.8 (Elementary move projection [41]) *If $\mu, \mu' \in \mathcal{M}(S)$ differ by one elementary move, then for any domain $Y \subset S$ with $\xi(Y) \neq 0$,*

$$d_Y(\mu, \mu') \leq 4.$$

3.3 Hierarchies

Thurston's classification theorem for surface homeomorphisms [52] gives a layered structure for homeomorphisms based upon studying subsurfaces which are preserved (perhaps under an iterate of the homeomorphism). In [41], Masur and Minsky refine and further elucidate a layered structure for the mapping class group. Using the marking complex and the complex of curves, they provide a way to compare how relatively complicated two given mapping class group elements are on any subsurface (not just ones that are eventually periodic!). Towards this end, they introduce an object called a *hierarchy* which ties all the subsurface comparisons together.

Definition 3.9 Let $Y \subset S$ with $\xi(Y) > 1$. A sequence of simplices $v_0, v_1, \ldots, v_n$ is called *tight* if:

(1) For each $0 \leq i < j \leq n$, and vertices $v_i' \subset v_i$ and $v_j' \subset v_j$,

$$d_Y(v_i', v_j') = |i - j|.$$

(2) For each $0 < i < n$ we have that v_i is the boundary of the subsurface filled by $v_{i-1} \cup v_{i+1}$.

When $\xi(Y) = 1$ we consider a sequence to be *tight* if and only if it is the vertex sequence of a geodesic.

When $\xi(Y) = -1$ we consider a sequence to be *tight* if and only if it is the vertex sequence of a geodesic where the set of endpoints on $\partial \widehat{Y}$ of arcs representing the vertices equals the set of endpoints of the first and last arc.

We often use the following decorated version of a tight sequence.

Definition 3.10 A *tight geodesic* g in $\mathcal{C}(Y)$ consists of a tight sequence $v_0, v_1 \ldots, v_n$ and a pair of markings $\mathbf{I} = \mathbf{I}(g)$ and $\mathbf{T} = \mathbf{T}(g)$ (called the *initial* and *terminal markings* for g) such that v_0 is a subset of the simplex base($\mathbf{I}$) and v_n is a subset of base($\mathbf{T}$).

The integer n is called the length of g. The *domain* (sometimes called the *support*) of g refers to the surface Y, written $D(g) = Y$.

Below we explain a relationship between tight geodesics occurring in different subsurfaces.

When $\mu \in \mathcal{M}(S)$ we write $\mu|_Y$ to denote the restriction of this marking to Y, by which we mean:

- If $\xi(Y) = -1$, then $\mu|_Y = \pi_Y(\mu)$.

- Otherwise, we let $\mu|_Y$ be the set of base curves which meet Y essentially, each taken with their associated transversal.

For a surface Y with $\xi(Y) \geq 1$ and a simplex $v \subset \mathcal{C}(Y)$, we say that X is a *component domain* of (Y, v) if either: X is a component of $Y \setminus v$ or X is an annulus whose core curve is a component of v. More generally, we say that a subsurface $X \subset Y$ is a *component domain of g* if for some simplex v_i in g, X is a component domain of $(D(g), v_i)$. Notice that v_i is uniquely determined by g and X.

Furthermore, when X is a component domain of g and $\xi(X) \neq 0$, define the *initial marking of X relative to g*:

$$\mathbf{I}(X, g) = \begin{cases} v_{i-1}|_X & v_i \text{ is not the first vertex (of } g) \\ \mathbf{I}(g)|_X & v_i \text{ is the first vertex} \end{cases}$$

Similarly define the *terminal marking of X relative to g* to be:

$$\mathbf{T}(X, g) = \begin{cases} v_{i+1}|_X & v_i \text{ is not the last vertex} \\ \mathbf{T}(g)|_X & v_i \text{ is the last vertex} \end{cases}$$

Observe that these are each markings, since ∂X is distance 1 in $\mathcal{C}(S)$ from each of $v_{i\pm1}$, or in the case where v_i is the first vertex, then ∂X is disjoint from base($\mathbf{I}(g)$) (similarly for the terminal markings).

When X is a component domain of g with $\mathbf{T}(X, g) \neq \varnothing$, then we say that X is *directly forward subordinate* to g, written $X \searrow^d g$. Similarly, when $\mathbf{I}(X, g) \neq \varnothing$ we say that X is *directly backward subordinate* to g, written $g \searrow^d X$.

The definition generalizes to geodesics as follows.

Definition 3.11 Let g and h be tight geodesics. We say that g is *directly forward subordinate* to h, written $g \searrow^d h$, when $D(g) \searrow^d h$ and $\mathbf{T}(g) = \mathbf{T}(D(g), h)$. Similarly, h is *directly backward subordinate* to g, written $g \swarrow^d h$, when $g \swarrow^d D(h)$ and $\mathbf{I}(h) = \mathbf{I}(D(h), g)$.

We write *forward subordinate*, or $\searrow$, to denote the transitive closure of $\searrow^d$; similarly we define $\swarrow$.

We can now define the main tool which was introduced in [41].

Definition 3.12 A *hierarchy (of geodesics) H*, on S is a collection of tight geodesics subject to the following constraints:

(1) There exists a tight geodesic whose support is S. This geodesic is called the *main geodesic* and is often denoted g_H. The initial and terminal markings of g_H are denoted $\mathbf{I}(H)$ and $\mathbf{T}(H)$.

(2) Whenever there exists a pair of tight geodesics $g, k \in H$ and a subsurface $Y \subset S$ such that $g \swarrow^d Y \searrow^d k$ then H contains a unique tight geodesic h with domain Y such that $g \swarrow^d h \searrow^d k$.

(3) For each geodesic $h \in H$ other than the main geodesic, there exists $g, k \in H$ so that $g \swarrow^d h \searrow^d k$.

Using an inductive argument, it is proved in [41] that given any two markings on a surface, there is a hierarchy which has initial marking one of them and terminal marking the other. The process begins with picking a base curve of the initial marking, one in the terminal marking, and a geodesic in $\mathcal{C}(S)$ between them (the main geodesic). Then, the second condition in the definition of a hierarchy forces certain proper subdomains to support geodesics: if there is a configuration $g \swarrow^d D \searrow^d k$ such that D does not support any geodesic h with $\mathbf{I}(h) = \mathbf{I}(D, g)$ and $\mathbf{T}(h) = \mathbf{T}(D, k)$, then we construct such a geodesic. When this geodesics is constructed, we can choose the initial vertex to be any element of base $\mathbf{I}(Y, g)$ (similarly for the terminal vertex). This process continues until enough geodesic are included in H so that the second and third conditions are satisfied. As the above suggests, there is in general not just one hierarchy, but many of them connecting any pair of markings. In our proof of Theorem 4.3 (Projection estimates) we will exploit this flexibility by building our hierarchy subject to certain additional constraints which we find useful.

When we consider several hierarchies at once, we use the notation $g_{H,Y}$ to denote the geodesic of H supported on Y, when this geodesic exists it is unique by Theorem 3.13.

For any domain $Y \subset S$ and hierarchy H, the *backward* and *forward sequences* are given respectively by

$$\Sigma_H^-(Y) = \{b \in H : Y \subseteq D(b) \text{ and } I(b)|_Y \neq \varnothing\}$$

and
$$\Sigma_H^+(Y) = \{f \in H : Y \subseteq D(b) \text{ and } T(f)|_Y \neq \varnothing\}.$$

The following theorem summarizes some results which are useful for making computations with hierarchies.

Theorem 3.13 (Structure of Sigma [41]) *Let H be a hierarchy, and Y any domain in its support.*

(1) *If $\Sigma_H^+(Y)$ is nonempty then it has the form of a sequence*

$$f_0 \searrow^d \cdots \searrow^d f_n = g_H,$$

where $n \geq 0$. Similarly, if $\Sigma_H^-(Y)$ is nonempty then it has the form of a sequence

$$g_H = b_m \nearrow^d \cdots \nearrow^d b_0,$$

where $m \geq 0$.

(2) *If $\Sigma_H^\pm(Y)$ are both nonempty and $\xi(Y) \neq 0$, then $b_0 = f_0$, and Y intersects every vertex of f_0 nontrivially.*

(3) *If Y is a component domain of a geodesic $k \in H$ and $\xi(Y) \neq 0$, then*

$$f \in \Sigma_H^+(Y) \iff Y \searrow f,$$

and similarly,
$$b \in \Sigma_H^-(Y) \iff b \nearrow Y.$$

If, furthermore, $\Sigma^\pm(Y)$ are both nonempty, then in fact Y is the support of the geodesic $b_0 = f_0$.

(4) *Geodesics in H are determined by their supports. That is, if $D(h) = D(h')$ for $h, h' \in H$ then $h = h'$.*

Given a hierarchy, the following provides a useful criterion for determining when a domain is the support of some geodesic in that hierarchy.

Lemma 3.14 (Large links [41]) *If Y is any domain in S and*

$$d_Y(I(H), T(H)) > M_2,$$

then Y is the support of a geodesic h in H, where M_2 only depends on the topological type of S.

The geodesics in a hierarchy admit a partial ordering, which generalizes both the linear ordering on vertices in a geodesic and the ordering coming from forward and backwards subordinacy. Below we recall the basic definitions and a few properties of this ordering.

Given a geodesic g in $\mathcal{C}_1(S)$ and a subsurface Y, define the *footprint* of Y in g, denoted $\phi_g(Y)$ to be the collection of vertices of g which are disjoint from Y. It is easy to see that the diameter of this set is at most 2 and with a little work it can be shown that the footprint is always an interval (of diameter at most 2). The proof that $\phi_g(Y)$ is interval uses the assumption that the geodesic is tight: this result is the only place where the tightness assumption on geodesics is utilized. That footprints form intervals is useful as it allows one to make the following definition.

Definition 3.15 For a pair of geodesics $g, k \in H$ we say *g precedes k in the time order*, or $g \prec_t k$, if there exists a geodesic $m \in H$ so that $D(g)$ and $D(k)$ are both subsets of $D(m)$ and

$$\max \phi_m(D(g)) < \min \phi_m(D(k)).$$

We call m the *comparison geodesic*.

Time ordering is a (strict) partial ordering on geodesics in a hierarchy, this is proven in [41]. It is worth remarking that when a pair of geodesics are time ordered, they are time ordered with respect to a unique comparison geodesic.

The following provides a way to use the time ordering to gain information about the hierarchy.

Lemma 3.16 (Order and projections [41]) *Let H be a hierarchy in S and $h, k \in H$ with $D(h) = Y$ and $D(k) = Z$. Suppose that $Y \cap Z \neq \varnothing$ and neither domain is contained in the other. Under these conditions, if $k \prec_t h$ then $d_Y(\partial Z, I(H)) \leq M_1 + 2$ and $d_Z(T(H), \partial Y) \leq M_1 + 2$. The constant M_1 only depends on the topological type of S.*

Since we will often use the above hypothesis, we introduce the terminology that a pair of subsurfaces Y and Z of S *overlap* when $Y \cap Z \neq \varnothing$ and neither domain is contained in the other. A application of this lemma which we shall often use is: given $h_1, h_2, k \in H$ with $D(h_i) = Z_i$ and $D(k) = Y$ and for which $h_i \prec_t k$, then $d_Y(\partial Z_1, \partial Z_2) \leq 2(M_1 + 2)$.

The next result provides a way to translate distance computations in the mapping class group to computations in curve complexes of subsurfaces.

Theorem 3.17 (Move distance and projections [41]) *There exists a constant $t(S)$ such that for each $\mu, v \in \mathcal{M}$ and any threshold $t > t(S)$, there exists $K(= K(t))$ and $C(= C(t))$ such that:*

$$\frac{1}{K}d_{\mathcal{M}}(\mu, v) - C \leq \sum_{\substack{Y \subseteq S \\ d_Y(\mu,v)>t}} d_Y(\mu, v) \leq Kd_{\mathcal{M}}(\mu, v) + C.$$

The importance of Theorem 3.17 can not be understated, as this theorem provides the crucial result that hierarchies give rise to quasi-geodesic paths in the mapping class group. We now indicate this construction.

Given a hierarchy H, [41] define a *slice*[3] in H to be a set τ of pairs (h, v), where $h \in H$ and v is a vertex of h, satisfying the following conditions:

(1) A geodesic h appears in at most one pair of τ.

(2) There is a distinguished pair (h_H, v_τ) appearing in τ, where h_H is the main geodesics of H.

(3) For every $(k, w) \in \tau$ other than (h_H, v_τ), $D(k)$ is a component domain of $(D(h), v)$ for some $(h, v) \in \tau$.

(4) Given $(h, v) \in \tau$, for every component domain Y of $(D(h), v)$ there is a pair $(k, w) \in \tau$ with $D(k) = Y$.

To any slice τ is associated a unique marking as follows. All the vertices which appear in geodesics of non-annular subsurfaces are distinct and disjoint, so these form a maximal simplex of $\mathcal{C}(S)$ which we take to be base(μ_τ). Then, by the forth condition, τ must contain a pair of the form (h_γ, t) for each $\gamma \in$ base(μ_τ). Letting t be the transversal for γ we obtain a complete marking, which we denote μ_τ. This marking may not be clean, but a lemma of [41] implies that a complete clean marking can be found in a uniformly bounded neighborhood; thus we assume $\mu_\tau \in \mathcal{M}(S)$.

The set of slices admits a partial ordering (relating to the time ordering discussed above), which then descends to provide a partial ordering on the markings μ_τ for a given hierarchy H. In [41] it is shown that slices nearby in the slice ordering yield markings which are close in $\mathcal{M}(S)$; combined with 3.17 these sequences of markings form quasi-geodesic paths in $\mathcal{M}(S)$, called *hierarchy paths*.

---

[3]This is called a *complete slice* in [41].

3.4 Teichmüller space and the pants complex

For a detailed reference on Teichmüller space, consult Imayoshi and Taniguchi [31].

Definition 3.18 For a fixed open topological surface S, the *Teichmüller space of S* is the space of equivalence classes of pairs (X, f), where X is a finite area hyperbolic surface and $f \colon S \to X$ is a homeomorphism. A pair (X_1, f_1) and (X_2, f_2) are considered equivalent if there exists an isometry $h \colon X_1 \to X_2$ such that $h \circ f_2$ is homotopic to f_1.

A topology is obtained on Teichmüller space by infimizing over the distortion of maps $h \colon X_1 \to X_2$. Topologically Teichmüller space is fairly easily understood, as the following classical result indicates:

Theorem 3.19 *Teichmüller space is homeomorphic to $\mathbb{R}^{6g-6+2p}$.*

There are several natural metrics on Teichmüller space and its metric structure is far less transparent than its topological structure. Often of interest is the natural complex structure which Teichmüller space carries. What will be important in the sequel is the *Weil–Petersson metric* on Teichmüller space, which is a Kähler metric with negative sectional curvature. We will not need the integral form of the definition, so we just mention that the metric is obtained by considering a natural identification between the complex cotangent space of Teichmüller space and the space of holomorphic quadratic differentials, then defining the Weil–Petersson metric to be the one dual to the L^2–inner product on the space of holomorphic quadratic differentials. See Wolpert [54] for a survey on the Weil–Petersson metric and its completion.

Definition 3.20 A *pair of pants* is a thrice punctured sphere. A *pants decomposition* of a surface $S_{g,p}$ is a maximal collection of pairwise non-parallel homotopy classes of simple closed curves; this decomposition obtains its name from the observation that such a curve system cuts $S_{g,p}$ into $2g + p - 2$ pairs of pants.

It is easy to verify that any pants decomposition consists of exactly $3g + p - 3$ disjoint simple closed curves on S and is thus a maximal simplex in $\mathcal{C}(S)$. These decompositions have long been useful in the study of the mapping class group and Teichmüller space.

Originally defined by Hatcher [28], the *pants complex of S*, denoted $\mathcal{P}(S)$, is a way of comparing all possible pants decompositions on a fixed surface. This complex consists of a vertex for every pants decomposition and an edge between each pair of decompositions which differ by an *elementary move*. Two pairs of pants $P = \{\gamma_1, \ldots, \gamma_{3g+p-3}\}$ and $P' = \{\gamma'_1, \ldots, \gamma'_{3g+p-3}\}$ differ by an *elementary move* if P and P' can be reindexed so that both:

(1) $\gamma_i = \gamma_i'$ for all $2 \le i \le 3g + p - 3$

(2) In the component T of $S \setminus \cup_{2 \le i \le 3g+p-3} \gamma_i$ which is not a thrice punctured sphere (T is necessarily either a once punctured torus or a 4 punctured sphere) we have (see Figure 3):
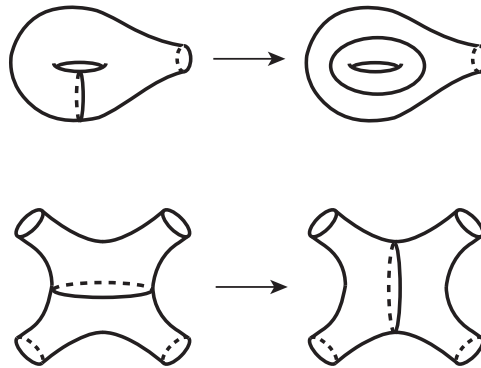
$$d_{\mathcal{C}(T)}(\gamma_1, \gamma_1') = 1.$$

Figure 3: Elementary moves in the pants complex

$\mathcal{P}(S)$ is metrized by giving each edge the metric of the Euclidean interval $[0, 1]$. Our interest in this space comes out of the following remarkable theorem of Brock:

Theorem 3.21 [14] *$\mathcal{P}(S)$ is quasi-isometric to the Teichmüller space of S with the Weil–Petersson metric.*

Remark 3.22 Noticing that a marking without transverse data is just a pants decomposition, in Section 8 of [41] it is remarked that all of the constructions in their paper using hierarchies to obtain results about the marking complex can be replaced with analogous theorems for the pants complex. This is done by replacing markings (and hierarchies, etc) by markings without transverse data (and hierarchies-without-annuli, etc). We will often implicitly use this result to note that results we prove from the mapping class group also hold in Teichmüller space (in particular, this allows one to obtain Teichmüller space analogues of the results in Sections 5 and 6). The following is the main result of this form which we will explicitly use.

Theorem 3.23 (Move distance and projections for $\mathcal{P}(S)$ [41]) *There exists a constant $t(S)$ such that for each $\mu, \nu \in \mathcal{P}$ (and any threshold $t > t(S)$) there exists $K(= K(t))$*

and $C(= C(t))$ *such that:*

$$\frac{1}{K} d_{\mathcal{P}}(\mu, \nu) - C \sum_{\substack{\text{non-annular } Y \subseteq S \\ d_Y(\mu,\nu) > t}} d_Y(\mu, \nu) \leq K d_{\mathcal{P}}(\mu, \nu) + C.$$

4 Projection estimates

In this section we address the following question: given a pair of subsurfaces X_1, X_2 of S, what is the image of $\mathcal{M}(S)$ under the map which projects markings into the product of curve complexes of these subsurfaces, ie, the image of $\mathcal{M}(S)$ as a subset of $\mathcal{C}(X_1) \times \mathcal{C}(X_2)$? We answer this in Theorems 4.3 and 4.4 using hierarchies in the curve complex as our main tool (see Section 3.1 for background). The geometric description underlying Theorem 4.4 can be generalized to arbitrary finite collections of subsurfaces by introducing an notion called a coarse tree-flat and using a delicate inductive step, this analysis is carried out in [3].

We begin this section with a pair of toy examples which illustrate the dichotomy of Theorem 4.3 (Projection estimates), but which can be proven without use of the full hierarchy machinery. In Section 4.2 we prove Theorem 4.3 and its geometric analogue Theorem 4.4.

4.1 A motivating example: the dichotomy

Here we give two computations on the genus three surface. We exhibit a dichotomy between the behavior of projection maps into disjoint and intersecting pairs of subsurfaces. These examples illustrate the geometric meaning of Theorem 4.3 and provide a useful warm up for its proof, as the arguments we use here are a particularly easy form of what in general is a technical argument involving hierarchies.

Example 4.1 (Disjoint subsurfaces) Let S be a closed genus three surface and let X and Y be subsurfaces of S which are each once punctured tori and which can be realized disjointly on S (Figure 4).

Let $\alpha \in \mathcal{C}(X)$ and $\beta \in \mathcal{C}(Y)$. Note that both these curves can also be considered as elements of $\mathcal{C}(S)$, then since $X \cap Y = \varnothing$ it follows that $d_{\mathcal{C}(S)}(\alpha, \beta) = 1$. Thus α and β are vertices of a common top dimensional simplex $\mu \subset \mathcal{C}(S)$. Taking a complete clean marking $\mu' \in \mathcal{M}(S)$ with base(μ') = μ, we have that the map

$$\mathcal{M}(S) \to \mathcal{C}(X) \times \mathcal{C}(Y)$$
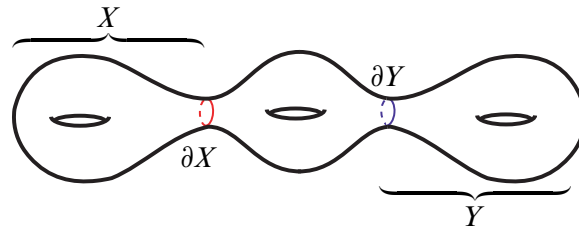
is onto.

Figure 4: S and two disjoint subsurfaces

The next example provides the more interesting half of the dichotomy.

Example 4.2 (Overlapping subsurfaces) Let S be the closed genus three surface and consider two subsurfaces X and Y which each have genus two and one puncture. Moreover, suppose that X and Y overlap in a twice punctured torus, Z. Letting μ be an element of $\mathcal{C}(S)$ which intersects both X and Y. We shall show that both $d_X(\mu, \partial Y)$ and $d_Y(\mu, \partial X)$ cannot simultaneously be large.
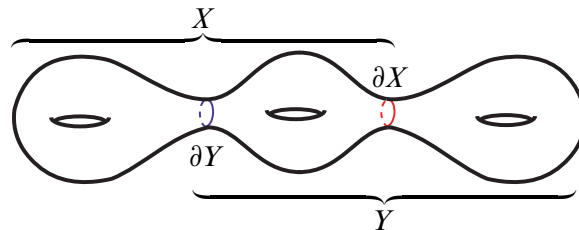

Figure 5: S and two overlapping subsurfaces

Let g be a geodesic in $\mathcal{C}_1(S)$ connecting μ to ∂X, with vertices $v_0, v_1, \ldots, v_n$ where $v_0 = \mu$ and $v_n = \partial X$. Since g is a geodesic in $\mathcal{C}(S)$, the curve v_{n-1} may be disjoint from X, but for all $k < n-1$ we have $v_k \cap \partial X \neq \varnothing$ and in particular, $\pi_X(v_k) \neq \varnothing$ and $\pi_Y(v_k) \neq \varnothing$. Since $S = X \cup Y$ we have either $v_{n-1} \cap Y \neq \varnothing$ or $v_{n-1} \subset X$.

If $v_{n-1} \cap Y \neq \varnothing$, then every vertex of g intersects Y non-trivially. Theorem 3.5 (Bounded geodesic image [41]) states that if g is a geodesic in $\mathcal{C}_1(S)$ with every vertex non-trivially intersecting a subsurface $Y \subset S$, then $\mathrm{diam}_Y(\pi_Y(g)) < K$ for a constant K depending only on the topological type of S. Accordingly, this theorem implies that we have $d_Y(\mu, \partial X) < K$ for a constant K depending only on the genus of S.

If $v_{n-1} \cap Y = \varnothing$, then it must be the case that $v_{n-1} \subset X$. In this case, we have $v_k \cap X \neq \varnothing$ for each $0 \leq k \leq n-1$. Applying the Bounded Geodesic Image Theorem tells us that $d_X(\mu, v_{n-1}) < K$. Since $v_{n-1} \cap Y = \varnothing$, we have in particular that

$v_{n-1} \cap \partial Y = \varnothing$ which implies that $d_X(v_{n-1}, \partial Y) \leq 3$. Thus we may conclude that $d_X(\mu, \partial Y) < K + 3$.

Putting the two cases together, this example shows that there exists a constant $C = K+3$, for which any $\mu \in \mathcal{C}(S)$ must satisfy either $d_X(\mu, \partial Y) < C$ or $d_Y(\mu, \partial X) < C$.

If we take μ to be one of the base curves for some complete clean marking ν of S, we have shown that for any marking $\nu \in \mathcal{M}(S)$ either

$$d_X(\nu, \partial Y) < C \text{ or } d_Y(\nu, \partial X) < C.$$

We consider this a toy example, since the only tool we needed was the Bounded geodesic image Theorem. Difficulties in generalizing this example to arbitrary overlapping subsurfaces include problems such as: the boundaries of these subsurfaces may fill S or the union of these subsurfaces may not be all of S—both properties that were crucial to the above argument. To deal with these problems we use a powerful technical tool called a hierarchy, in which one considers not just a geodesic in the curve complex of S, but a family of geodesics each in the curve complex of some subsurface of S. The tools introduced by Masur and Minsky for calculating with these allow one to give the argument in the next section which is obtained by bootstrapping the key idea used in the above example.

4.2 Projection estimates theorem

Theorem 4.3 (Projection estimates) *Let Y and Z be two overlapping surfaces in S, with $\xi(Y) \neq 0 \neq \xi(Z)$, then for any $\mu \in \mathcal{M}(S)$:*

$$d_Y(\partial Z, \mu) > M \implies d_Z(\partial Y, \mu) \leq M,$$

for a constant M depending only on the topological type of S.

Proof Let M_1 and M_2 be given by Lemmas 3.14 and 3.16 (Large link Lemma and Order and projections Lemma [41]). Define:

$$M = \max\{M_1 + 2, M_2 + 3\}.$$

We now fix a point $\mu \in \mathcal{M}$ for which $d_Y(\mu, \partial Z) > M$ and will show that this implies a bound on $d_Z(\mu, \partial Y)$.

Consider a hierarchy H with $\text{base}(I(H)) \supset \pi_Z(\mu) \cup \partial Z$ and $T(H) = \mu$ (when Z is an annulus take $\text{base}(I(H)) = \partial Z$ with transversal $\pi_Z(\mu)$). Furthermore, we assume H is built subject to the following two constraints. (See Section 3.3 for a discussion of the choices involved when building a hierarchy.)

- Choose the initial vertex of g_H to be an element of ∂Z which intersects Y. Call this vertex v_0.
- When adding a geodesic to H with a given initial marking $I(Y, g)$, if $v_0 \in \text{base}(I(Y, g))$ choose this as the initial vertex for the geodesic. If $v_0 \notin \text{base}(I(Y, g))$, then if any elements of ∂Z are in $\text{base}(I(Y, g))$ choose one of these to be the initial vertex for the geodesic.

The proof is now broken into four steps. We show there exists a geodesic $k \in H$ with domain Z and a geodesic $h \in H$ with domain Y. Then we show $k \prec_t h$, from which the theorem follows as a consequence of Lemma 3.16 (Order and projections [41]).

Step (i) *There exists a geodesic k with domain Z.*

We start by considering the forward and backward sequences, $\Sigma_H^+(Z)$ and $\Sigma_H^-(Z)$, as defined in the discussion preceding Theorem 3.13 (Structure of Sigma [41]).

Since μ is complete we have $\pi_Z(\mu) \neq \varnothing$; thus both the initial and terminal markings restrict to give nontrivial markings on Z (when Z is an annulus, we are using that $I(H)$ has a transverse curve). This implies for any subsurface $Q \supseteq Z$ that $\Sigma_H^-(Q)$ and $\Sigma_H^+(Q)$ both contain g_H and thus in particular are nonempty. Theorem 3.13 shows that when Q is a component domain of a geodesic in H and both $\Sigma_H^+(Q)$ and $\Sigma_H^-(Q)$ are non-empty, then Q must be the support of a geodesic in H.

Since the first vertex of g_H is $v_0 \in \partial Z$, we know it has a component domain $Q_1 \subsetneq S$ which contains Z. By the above observation, we know that $\Sigma_H^+(Q_1)$ and $\Sigma_H^-(Q_1)$ are both non-empty, and thus Q_1 supports a geodesic which we call k_1. If $Q_1 = Z$ we have produced a geodesic supported on Z and are done, otherwise since $Q_1 \supsetneq Z$ we can choose an element of ∂Z as the first vertex of the geodesic supported in Q_1.

Starting from the base case $Q_0 = S$ and $k_0 = g_H$, the above argument produces a sequence of properly nesting subsurfaces $S = Q_0 \supsetneq Q_1 \supsetneq \ldots \supsetneq Q_n = Z$ where for each $i > 0$ the subsurface Q_i is a component domain of a geodesic $k_{i-1} \in H$ supported in Q_{i-1} and each of the k_i has an element of ∂Z as an initial vertex.

Since S is a surface of finite type, the above sequence of nested surfaces must terminate with Z after finitely many steps.

Thus Z is a component domain of a geodesic in S and supports a geodesic which we call k.

Step (ii) *There exists a geodesic h with domain Y.*

By hypothesis we have $d_Y(\partial Z, \mu) > M_2 + 3$ and thus:

$$d_Y(I(H), T(H)) = d_Y(\pi_Z(\mu) \cup \partial Z, \mu) > M_2.$$

Together with Lemma 3.14 (Large link Lemma [41]), this implies Y is the support of a geodesic h in H.

Step (iii) $k \prec_t h$

In this step we will prove that the geodesic k precedes h in the time ordering on geodesics in H.

Since $D(k) = Z$ and $D(h) = Y$ and each are contained in $D(g_H) = S$: if $\max \phi_{g_H}(Z) < \min \phi_{g_H}(Y)$, then $k \prec_t h$ which is what we wanted to prove. In the general case, we provide an inductive procedure to show that these geodesics have the desired time ordering.

We refer to the ordered vertices of g as $v_i(g)$. Recall that by the first part of the constraint we have $v_0(g_H) = v_0$, where v_0 was chosen to satisfy $v_0 \in \partial Z$ and $v_0 \cap Y \neq \varnothing$. Also, since g_H is a tight geodesic, $v_1(g_H) = \partial F(v_0(g_H), v_2(g_H))$ where $F(\alpha, \beta)$ denotes the surface filled by α and β.

Summarizing, we have:

- $Y \cap v_0 \neq \varnothing$, and thus $v_0 \notin \phi_{g_H}(Y)$.
- Z is contained in a component domain of v_0, thus $v_0 \in \phi_{g_H}(Z)$.
- Since g_H is a geodesic, we know $v_2(g_H)$ must intersect $v_0 (\in \partial Z)$. Since the diameter of a footprint is at most 2, it now follows that $\max \phi_{g_H}(Z) \leq v_1(g_H)$.

Together these imply:

$$\max \phi_{g_H}(Z) \leq v_1(g_H) \text{ and } v_1(g_H) \leq \min \phi_{g_H}(Y).$$

Now there are two mutually exclusive cases to consider:

(1) $v_1(g_H) \cap Z \neq \varnothing$ or $v_1(g_H) \cap Y \neq \varnothing$

(2) $v_1(g_H) \cap Z = \varnothing = v_1(g_H) \cap Y$.

In the first case, depending on which of the two sets is non-empty, we obtain $v_1(g_H) \notin \phi_{g_H}(Z)$ or $v_1(g_H) \notin \phi_{g_H}(Y)$, respectively. Either of these imply $\max \phi_{g_H}(Z) < \min \phi_{g_H}(Y)$, proving that in case (1) we have $k \prec_t h$.

Observing that in $S_{1,1}$ and $S_{0,4}$ the footprint of a domain must consist of at most one vertex, we have $v_1(g_H) \notin \phi_{g_H}(Z)$ and thus when S is either of these surfaces we are

in case (1) and thus $k \prec_t h$. As the annulus does not contain any pair of surfaces which overlap we may now assume for the rest of the proof that $\xi(S) > 1$. In particular, for the remainder of the argument we need not consider the sporadic cases where $\mathcal{C}(S)$ has a special definition and thus we can assume two distinct homotopy classes of curves have distance one in the curve complex if and only if they can be realized disjointly.

For the remainder of this step we assume that $v_1(g_H) \in \phi_{g_H}(Z)$ and $v_1(g_H) \in \phi_{g_H}(Y)$ and will prove that we still obtain $k \prec_t h$. Since $v_1(g_H) \cap Z = \varnothing$ there is a component of $S \setminus v_1(g_H)$ which contains Z and since $Z \cap Y \neq \varnothing$ the same component contains Y as well—we will call this component W_1. Since W_1 is a component domain of H which intersects both $I(H)$ and $T(H)$ (since $W_1 \supset Z$), Theorem 3.13 (Structure of Sigma) implies that it supports a geodesic l_1.

Since W_1 is a component domain of $v_1(g_H)$, we have $v_0 \in I(W_1, g_H)$. Thus, by our convention for choosing geodesics in H, we choose $v_0(l_1) = v_0$. As before we have $v_0(l_1) \in \phi_{l_1}(Z)$, $v_2(l_1) \notin \phi_{l_1}(Z)$, and $v_0(l_1) \notin \phi_{l_1}(Y)$. If $v_1(l_1)$ is in both $\phi_{l_1}(Z)$ and $\phi_{l_1}(Y)$ then we again restrict ourselves to the appropriate component W_2 of $W_1 \setminus v_1(l_1)$ and repeat the argument with W_2 and l_2. This gives a properly nested collection of subsurfaces each of which contains $Y \cup Z$, so the process terminates in a finite number of steps to produce a geodesic l_n with domain $W_n \supseteq Y \cup Z$ and which satisfies $v_0(l_n) = v_0 \in \phi_{l_n}(Z)$, $v_2(l_n) \notin \phi_{l_n}(Z)$, $v_0(l_n) \notin \phi_{l_n}(Y)$, and either $v_1(l_n) \notin \phi_{l_n}(Z)$ or $v_1(l_n) \notin \phi_{l_n}(Y)$. Hence $\max \phi_{l_n}(Z) < \min \phi_{l_n}(Y)$, and we have $k \prec_t h$ with comparison geodesic l_n.

Step (iv) *Conclusion*

We have now produced a hierarchy H with geodesics k and h with $D(h) = Y$, $D(k) = Z$, and $k \prec_t h$. Thus Lemma 3.16 (Order and projections [41]) implies that $d_Z(\partial Y, T(H)) \leq M_1 + 2$.

Since $T(H) = \mu$, this implies $d_Z(\partial Y, \mu) \leq M_1 + 2$ which is exactly what we wanted to show. $\qquad \square$

It would be interesting to have a more elementary proof of Theorem 4.3, although the given argument is useful in that it provides the reader an introduction to the tools we need for the later parts of this work. C Leininger has suggested to me ideas for a more constructive proof.

4.2.1 Geometry of projection estimates The main case of the following result is an immediate corollary of Theorem 4.3 (Projection estimates); we give this restatement because it emphasizes the geometry underlying that result.

Theorem 4.4 (Projection estimates, geometric version) *For any distinct subsurfaces Y and Z of S, exactly one of the following holds for the map*

$$\pi_Y \times \pi_Z \colon \mathcal{M}(S) \to \mathcal{C}(Y) \times \mathcal{C}(Z)$$

(1) *$Y \cap Z = \varnothing$ in which case the map is onto.*

(2) *One of the surfaces is contained inside the other, say $Z \supset Y$. Then the image is contained in a radius 3 neighborhood of the set*

$$\mathcal{C}(Y) \times \pi_Z(\partial Y) \bigcup \mathrm{Graph}(\pi_{Z \to Y}),$$

where $\pi_{Z \to Y}$ is the projection map from $\mathcal{C}(Z) \setminus B_1(\partial Y)$ to $\mathcal{C}(Y)$. Furthermore, $\pi_Y \times \pi_Z(\mathcal{M}(S))$ is 6–dense in this set.

(3) *Y and Z overlap, in which case the image is contained in a radius $M = \max\{M_1 + 2, M_2\}$ neighborhood of the set*

$$\pi_Y(\partial Z) \times \mathcal{C}(Z) \bigcup \mathcal{C}(Y) \times \pi_Z(\partial Y).$$

The constant M depends only on the topological type of S. Furthermore, $\pi_Y \times \pi_Z(\mathcal{M}(S))$ is $6 \cdot M$–dense in this set.

Proof

Case 1 $Y \cap Z = \varnothing$

Any pair consisting of a curve in $\mathcal{C}(Z)$ and a curve in $\mathcal{C}(Y)$ can be completed to a complete marking on S so the map is onto.

Case 2 $Y \subset Z$

For any $\mu \in \mathcal{M}$ we have that $d_Z(\pi'_Z(\mu), \pi_Z(\mu)) \leq 1$ and thus when we restrict these arcs (curves) to Y we get that $d_Y(\pi'_Y \pi'_Z(\mu), \pi'_Y \pi_Z(\mu)) \leq 1$. Applying Lemma 3.3 (Lipschitz projection [41]), we have $d_Y(\phi \pi'_Y \pi'_Z(\mu), \phi \pi'_Y \pi_Z(\mu)) \leq 3$. Thus we have $d_Y(\mu, \pi_{Z \to Y} \circ \pi_Z(\mu)) \leq 3$.

This proves that if $Y \subset Z$ then $\pi_Y \times \pi_Z(\mathcal{M})$ is contained in a radius 3 neighborhood of $\mathcal{C}(Y) \times \pi_Z(\partial Y) \bigcup \mathrm{Graph}(\pi_{Z \to Y})$.

We now prove the density statement. Let $(\gamma, \delta) \in \mathrm{Graph}(\pi_{Z \to Y}) \subset \mathcal{C}(Y) \times \mathcal{C}(Z)$. Then take any $\mu \in \mathcal{M}(S)$ which has $\delta \in \mathrm{base}(\mu)$. By Lemma 3.3 we have $\pi_Z(\mu) \subset B_3(\delta)$. By hypothesis we have $\gamma \in \pi_Y(\delta)$, so Lemma 3.3 also shows $\pi_Y(\mu) \subset B_3(\gamma)$. Thus we have $(\gamma, \delta) \subset B_6(\pi_Y(\mu) \times \pi_Z(\mu))$. Similarly, given any $\gamma \in \mathcal{C}(Y)$, since $Y \subset Z$ we have γ and ∂Y are disjoint. Thus if $\mu \in \mathcal{M}(S)$ is chosen so that both γ and ∂Y

are in base(μ), we have $\pi_Z(\mu) \subset B_3(\gamma)$ and $\pi_Y(\mu) \subset B_3(\gamma)$. Thus $(\gamma, \pi_Z(\partial Y)) \subset B_6(\pi_Y(\mu) \times \pi_Z(\mu))$. Thus we have shown that $\mathcal{C}(Y) \times \pi_Z(\partial Y) \bigcup \text{Graph}(\pi_{Z \to Y}) \subset B_6(\pi_Y \times \pi_Z(\mathcal{M}))$.

Case 3 $Y \cap Z$, $Y \nsubseteq Z$, and $Z \nsubseteq Y$

If $\mu \in \mathcal{M}(S)$ projects to $\mathcal{C}(Y)$ with $d_Y(\partial Z, \mu) > M$ then Theorem 4.3 (Projection estimates) implies that $d_Z(\partial Y, \mu) \leq M$. Thus the image of the map $\mathcal{M}(S) \to \mathcal{C}(Y) \times \mathcal{C}(Z)$ is contained in the union of the radius M neighborhood of $\pi_Y(\partial Z) \times \mathcal{C}(Z)$ with the radius M neighborhood of $\mathcal{C}(Y) \times \pi_Z(\partial Y)$ as claimed.

Given any $\gamma \in \mathcal{C}(Y)$, Lemma 3.3 implies that $\text{diam}_Z(\pi_Z(\gamma), \pi_Z(\partial Y)) \leq 3$. If $\mu \in \mathcal{M}(S)$ is chosen so that both γ and ∂Y are in base(μ), then we have $\pi_Z(\mu) \subset B_3(\pi_Z(\partial Y))$ and $\pi_Y(\mu) \subset B_3(\gamma)$. Thus $(\gamma, \pi_Z(\partial Y)) \subset B_6(\pi_Y(\mu) \times \pi_Z(\mu))$. Thus we have shown that $\mathcal{C}(Y) \times \pi_Z(\partial Y) \subset B_6(\pi_Y \times \pi_Z(\mathcal{M}))$. The identical argument with the roles of Y and Z reversed completes the proof of the density statement. $\square$

Remark 4.5 In the above proposition we see that except in the case where $Y \cap Z = \varnothing$, there are uniform bounds for which the image of the map from $\mathcal{C}(S)$ to $\mathcal{C}(Y) \times \mathcal{C}(Z)$ lies in a neighborhood of the δ–hyperbolic space formed from the complexes of curves for Y and Z "joined together." The two complexes are glued together along a bounded diameter set in the non-nesting case. In the case of nesting $\mathcal{C}(Y)$ is glued to the link of a point in $\mathcal{C}(Z)$; essentially this is done by taking a "blow-up" of $\mathcal{C}(Z)$ at the point $\partial Y \in \mathcal{C}(Z)$. δ–hyperbolicity in the nesting case follows from a more delicate analysis, see [3].

Theorem 4.4 (Projection estimates, geometric version) is summarized in Figure 6. For a (slightly) more accurate picture of the behavior in the nesting case, see also Figure 1.
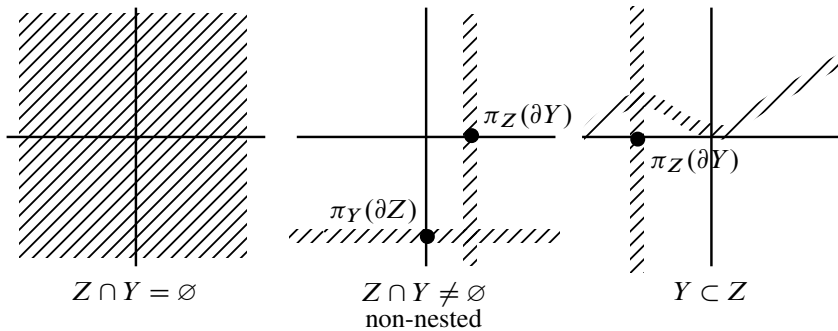

Figure 6: A cartoon of the possible images of projections from complete markings into $\mathcal{C}(Z) \times \mathcal{C}(Y)$

5 Hyperbolicity of the mapping class group and Teichmüller space in the cases of low complexity

In this section we give new proofs of the following theorems.

Theorem 5.1 $\mathcal{P}(S_{1,2})$ *and* $\mathcal{P}(S_{0,5})$ *are each* δ*–hyperbolic.*

Theorem 5.2 $\mathcal{M}(S_{1,1})$ *and* $\mathcal{M}(S_{0,4})$ *are each* δ*–hyperbolic.*

Theorem 5.2 is classical, since $\mathcal{M}(S_{1,1})$ is isomorphic to $SL_2(\mathbb{Z})$ and $\mathcal{M}(S_{0,4})$ to a finite subgroup of $SL_2(\mathbb{Z})$, hence both group are virtually free and thus easily seen to be δ–hyperbolic. Theorem 5.1 is a much deeper theorem, and was originally proved by Brock–Farb in [16] (see also [2]). Although the Weil–Petersson metric has been known for some time to have negative curvature, it was only recently discovered in [30] that the sectional curvature is not bounded away from zero, even for the case of $\mathcal{P}(S_{1,2})$ and $\mathcal{P}(S_{0,5})$, thereby prohibiting a proof of Theorem 5.1 by a comparison geometry argument.

We find especially tantalizing the phenomenon that these two results about two very different spaces can be proven simultaneously. This can be done since the property of low complexity which we use is that any two proper subsurfaces (with nontrivial curve complex) must overlap: a property of $S_{1,1}$ and $S_{0,4}$ which is also true for $S_{1,2}$ and $S_{0,5}$ when one adds the assumption that the subsurfaces are not annuli, a natural assumption when one is considering pants decompositions instead of markings (cf. Theorem 3.23). Thus the proof of Theorem 5.1 is obtained by rereading the proof for Theorem 5.2 reading "marking" as "pants decomposition," "hierarchy" as "hierarchy without annuli," and $\mathcal{M}(S)$ as $\mathcal{P}(S)$. The core arguments in this section will be revisited to prove Theorem 6.5, a result which holds for surfaces of arbitrary complexity.

Our method of proof is to show that for every hierarchy path there exists an "almost locally constant" map sending the marking (or pants) complex to the hierarchy path. That this implies hyperbolicity is a consequence of a general result from [42]; we also provide a new proof of this implication using asymptotic cones (Theorem 5.7).

5.1 A projection from markings to hierarchies

Fix a surface S, a pair of complete clean markings I and T, and a hierarchy H connecting them. By a hierarchy path, we mean a "resolution" of slices of the hierarchy into a sequence of markings separated by elementary moves; this is done by taking the slices of H and interpolating between them to get a path. The choice of a hierarchy

path is not canonical, but they are each (K, C)–quasi-geodesics with the quasi-isometry constants depending only on the topological type of the surface. (See [41], especially Section 5 and the Efficiency of hierarchies Theorem of Section 6.)

We say that Y *appears as a large domain in* H if Y is in the set

$$\mathcal{G} = \{Y \subseteq S : d_Y(I(H), T(H)) > 6M + 4\delta'\}.$$

Where $M = \max\{M_1 + 2, M_2 + 3\}$ is the constant coming from Theorem 4.3 (Projection estimates), note in particular that this constant is sufficiently large to satisfy the hypothesis of Lemma 3.14 (Large link [41]) and $\delta' = 4\delta + 5$ where δ denotes the maximum of the δ–hyperbolicity constants for $\mathcal{C}(Y)$ where $Y \subseteq S$. Alternatively, when dealing with the pants complex we consider:

$$\mathcal{G}' = \{Y \in \mathcal{G} : Y \text{ is not an annulus}\}.$$

Letting $g_{H,Y}$ denote the geodesic segment supported on Y in the hierarchy H, for each $Y \in \mathcal{G}$ we have a map $\mathcal{M}(S) \xrightarrow{\pi_Y} \mathcal{C}(Y) \xrightarrow{r_Y} g_{H,Y}$ where r_Y is the closest point(s) projection from $\mathcal{C}(Y)$ to $g_{H,Y}$. We denote the composition

$$p_Y = r_Y \circ \pi_Y \colon \mathcal{M}(S) \to g_{H,Y},$$

or just p when the surface Y is understood. Notice that δ–hyperbolicity of $\mathcal{C}(Y)$ (via Lemma 2.3) combined with the fact that $\operatorname{diam}_Y(\pi_Y(\mu)) \le 3$ (Lemma 3.3) imply that for any $\mu \in \mathcal{M}(S)$, the set $p_Y(\mu)$ has diameter at most δ'. Accordingly, it follows that maps p_Y are coarsely distance decreasing in the sense that:

$$(5\text{--}1) \qquad d_{\mathcal{C}(Y)}(p_Y(\mu), p_Y(\nu)) \le d_{\mathcal{C}(Y)}(\pi_Y(\mu), \pi_Y(\nu)) + 2\delta'.$$

Define $\qquad \mathcal{L}(\mu) = \{Y \in \mathcal{G} : d_Y(p_Y(\mu), T(Y)) < 3M + 2\delta'\}$

and $\qquad \mathcal{R}(\mu) = \{Y \in \mathcal{G} : d_Y(p_Y(\mu), I(Y)) < 3M + 2\delta'\}.$

Usually the choice of marking μ will be fixed and we will drop the μ from the notation, just writing $\mathcal{L}$ and $\mathcal{R}$.

The next lemma summarizes some basic facts we will need about these sets.

Lemma 5.3   (1) $\mathcal{L} \cap \mathcal{R} = \varnothing$

   (2)   Let $Y, Z \in \mathcal{G}$. If $Y \prec_t Z$ and Y and Z overlap, then either $Y \in \mathcal{L}$ or $Z \in \mathcal{R}$.

   (3)   If S is $S_{1,1}$ or $S_{0,4}$ then the set $\mathcal{G} \setminus (\mathcal{L} \cup \mathcal{R})$ consists of at most one surface.

   (4)   If S is $S_{1,2}$ or $S_{0,5}$ then $\mathcal{G}' \setminus (\mathcal{L} \cup \mathcal{R})$ consists of at most one surface.

Proof

(1) If $Y \in \mathcal{L}$ then $d_Y(p(\mu), T) < 3M$. If Y is also in $\mathcal{R}$ then $d_Y(p(\mu), I) < 3M$. But together these imply that the total distance H travels through Y is less than $6M$ contradicting our assumption that Y is a large domain.

(2) Suppose not. Then we have $Y \prec_t Z$, $Y \notin \mathcal{L}$, and $Z \notin \mathcal{R}$.

Since Y and Z are large domains with $Y \prec_t Z$, by [Lemma 3.16](#) we have:

$$d_Y(\partial Z, T(H)) \leq M_1 + 2 \text{ and } d_Z(\partial Y, I(H)) \leq M_1 + 2.$$

$Y \notin \mathcal{L}$ implies that

$$d_Y(p(\mu), T(Y)) \geq 3M + 2\delta'.$$

Similarly, $Z \notin \mathcal{R}$ implies that

$$d_Z(p(\mu), I(Z)) \geq 3M + 2\delta'.$$

Putting these facts together with the result from [41] that for any domain $d_D(I(D), I(H)) \leq M_1$, yields the two inequalities:

$$d_Y(p(\mu), \partial Z) > 3M + 2\delta' - M - (M_1 + 2) > M + 2\delta'$$

and $\qquad d_Z(p(\mu), \partial Y) > 3M + 2\delta' - M - (M_1 + 2) > M + 2\delta'.$

The fact that p_Y and p_Z are coarsely distance decreasing implies, via Equation [(5–1)](#), that both $d_Y(\mu, \partial Z) > d_Y(p(\mu), \partial Z) - 2\delta' > M$ and $d_Z(\mu, \partial Y) > d_Z(p(\mu), \partial Y) - 2\delta' > M$. But this is impossible, since together these inequalities contradict [Theorem 4.3](#) (Projection estimates).

(3) This follows from part 2, since in each of these surfaces any two subsurfaces overlap and thus are time ordered by Lemma 4.18 of [41].

(4) Again this follows from part 2, since here any two non-annular surfaces must overlap. $\qquad \square$

Parts 3 and 4 of the above proposition suggests the following construction, which we give for the marking complexes for $S_{1,1}$ and $S_{0,4}$ and also for the pants complexes of $S_{1,2}$ and $S_{0,5}$ (aspects of this construction generalize to work on arbitrary surfaces, we discuss this further in [Section 6](#)). Given any pair of markings $I, T \in \mathcal{M}(S)$ and a hierarchy H connecting them let $\widetilde{H}$ denote the set of markings associated to slices of H. Below we construct a map $\widehat{\Phi}$ which maps elements of $\mathcal{M}(S)$ to (uniformly) bounded diameter subsets of $\widetilde{H}$ (in the cases we are considering, $\widetilde{H}$ is metrized by a linear time ordering on slices). $\widetilde{H}$ be embedded into $\mathcal{M}(S)$ via a quasi-isometric

embedding which can be extended by resolving the set of slices into a path (referred to as a *hierarchy path*), the map $\widehat{\Phi}$ induces a map to (uniformly) bounded diameter subsets of any hierarchy path from I to T. The map $\widehat{\Phi}$ is defined to be the identity on markings in $\widetilde{H}$; the following defines the map for $\mu \in \mathcal{M}(S) \setminus \widetilde{H}$:

(1) If $\mathcal{G} \setminus (\mathcal{L} \cup \mathcal{R}) = \{A\} \neq \varnothing$, then define $\widehat{\Phi}(\mu) = (A, p_A(\mu))$.

(2) If $\mathcal{G} \setminus (\mathcal{L} \cup \mathcal{R}) = \varnothing$, then consider

$$\Lambda = \{v \in g_H : d_{\mathcal{C}(S)}(v, \mu) \leq d_{\mathcal{C}(S)}(w, \mu) \text{ for any } w \in g_H\},$$

ie, the set of points on the main geodesic g_H which are closest to μ, namely $r_S(\mu)$. $\widehat{\Phi}(\mu)$ is defined to be the following sets of markings.

(a) If any of the surfaces in $\mathcal{L}$ are component domains of g_H at a vertex in the set Λ, then denote by L the rightmost (the last to appear with respect to the time ordering) domain in $\mathcal{L}$, and define $\widehat{\Phi}(\mu)$ to contain each (L, v) where v is any vertex of $g_{H,L}$ within $3M + 3\delta'$ of the terminal marking T.

(b) If a component domain corresponding to a vertex of g_H in the set Λ is an element of $\mathcal{R}$ then choose the leftmost element of this set (call this R). Define $\widehat{\Phi}(\mu)$ to contain each of the markings (R, v) where v is any vertex of $g_{H,R}$ within $3M + 3\delta'$ of the initial marking I.

(c) For each domain D which is a component domain of a vertex in Λ and which is not in $\mathcal{G}$, if D supports a geodesic and it is time ordered after the L geodesic and before the R geodesic then define $\widehat{\Phi}(\mu)$ to contain each (D, v) where v is any vertex in $g_{H,D}$.

Before proceeding, we mention why $(A, p_A(\mu))$, (L, v), and (R, v), although not necessarily complete clean markings, give rise to such (in the text we abuse notation and drop this correspondence from the discussion, thereby simply referring to $(A, p_A(\mu))$, etc, as complete clean markings). In each case these consist of a proper subsurface X and a point in $\mathcal{C}(X)$ (we unify the argument that these give markings by writing $(X, p_X(\mu))$). For $S_{1,1}$ and $S_{0,4}$, a point of $\mathcal{M}(S)$ consist of a curve in $\mathcal{C}(S)$ and a transversal to that curve, so it is clear that our prescription above indeed describes a marking since the only nontrivial proper subsurfaces are annuli so $(X, p_X(\mu))$ refers to the marking $(\partial X, p_X(\mu))$. The case of $\mathcal{P}(S)$ for the surfaces $S_{1,2}$ and $S_{0,5}$ requires (only slightly) more justification. In these cases an element of $\mathcal{P}(S)$ is a pair of disjoint curves. Recall that the only subsurfaces of S with nontrivial curve complex are once punctured tori and four punctured spheres (we ignore annuli when dealing with the pants complex). Thus $p_X(\mu) \in \mathcal{C}(X)$ is also an element of $\mathcal{C}(S)$. Furthermore, $p_X(\mu)$ and ∂X have distance 1 in $\mathcal{C}(S)$, so $(X, p_X(\mu))$ is taken to refer to the pants decomposition $(\partial X, p_X(\mu))$.

Lemma 5.4 *There is a uniform bound D (depending only on the topological type of S), so that for each $\mu \in \mathcal{M}$ the set $\widehat{\Phi}(\mu)$ has diameter less than D.*

Proof First note that the set of points on the main geodesic g_H which are closest to μ has diameter at most δ' by Lemma 2.3. When $\mathcal{G} \setminus (\mathcal{L} \cup \mathcal{R}) \neq \varnothing$ there is only one such surface, as proven in Lemma 5.3. In this case it follows from the definition of Φ and δ–hyperbolicity of $\mathcal{C}(A)$ that $\Phi(\mu)$ is a subset of $\widetilde{H}$ of diameter at most δ'.

Note that under Φ the marking μ can not project to anything time ordered before the rightmost element $L \in \mathcal{L}$ as then Lemma 3.16 would force $p_L(\mu)$ to lie near the initial marking of L and then we would have either $L \notin \mathcal{G}$ or $L \in \mathcal{R}$, either way contradicting Lemma 5.3 which proves that every surface in $\mathcal{L}$ is time ordered before every surface of $\mathcal{R}$; a similar argument gives the analogous result for $\mathcal{R}$.

So now $\Psi(\mu)$ consists of the rightmost element of $\mathcal{L}$ which we call L, the leftmost element of $\mathcal{R}$ which we call R, and all the rest of the small domains supporting geodesics which are time ordered between L and R (of which there are at most δ'). Φ was defined to be the union of the three set (any of which are possibly empty):

(1) If $L \in \Lambda$ then (L, v) where v is any vertex of $g_{H,L}$ within $3M + 3\delta'$ of the terminal marking T.

(2) If $R \in \Lambda$ then (R, v) where v is any vertex of $g_{H,R}$ within $3M + 3\delta'$ of the initial marking I.

(3) If $D \in \Lambda$, $D \notin \mathcal{G}$, and $L \prec_t D \prec_t R$ then (D, v) where V is any vertex of $g_{H,D}$.

First, note that the diameter of the elements in item 1 is at most $B(3M + 3\delta')$, where B is the Lipschitz constant for the map from hierarchy slices to paths in the marking complex. (Similarly for the markings in item 2.) Since $\mathcal{C}(S)$ is δ–hyperbolic $\mathrm{diam}_{\mathcal{C}(S)}(\Lambda) < \delta'$, and thus if both L and R are in Λ then $d_{\mathcal{C}(S)}(L, R) < \delta'$. Also note that for any $D \in \Lambda$ as described in item 3 we have that the diameter in $\mathcal{M}(S)$ of the set of (D, v) is less than $B(6M + 4\delta')$ (since $6M + 4\delta'$ is the threshold for being in $\mathcal{G}$). Again, since the diameter in $\mathcal{C}(S)$ is bounded by δ' we then see that the set of all markings in item three has diameter in the marking complex bounded by $\delta' \cdot B(6M + 4\delta')$.

So now it follows that although $\Phi(\mu)$ consist of many slices:

$$\mathrm{diam}_{\mathcal{M}(S)}(\Phi(\mu)) < 2B(3M + 3\delta' + \delta'(6M + 4\delta')). \qquad \square$$

Note that since the map from slices to hierarchy paths is Lipschitz, the above lemma yields a map which sends points of $\mathcal{M}$ to uniformly bounded diameter subsets of the hierarchy path, where the constants are independent of the hierarchy path.

In the next subsection we will show that these maps to hierarchy paths are "almost locally constant" off the hierarchy path, in the sense that there are uniform constants, so that large diameter balls around markings far from a hierarchy path are sent to small diameter subsets of the hierarchy path.

5.2 Coarsely contracting projections

We now describe the tool we use for proving δ–hyperbolicity. It is a generalization to our context of Morse's Lemma on stability of quasi-geodesics in hyperbolic space.

Definition 5.5 In a space X, we say a family of paths $\mathcal{H}$ is *transitive* if every pair of points in X can be connected by a path in $\mathcal{H}$. We say $\mathcal{H}$ has the *coarsely contracting property* when $\mathcal{H}$ is a transitive family of paths in X with the property that for every $H \in \mathcal{H}$ there exists a map $\widehat{\Phi}_H \colon X \to H$ and constants b and c such that each for each $\mu, \mu' \in X$ satisfying $d_X(\mu, \mu') < b \cdot d_X(\mu, \widehat{\Phi}_H(\mu))$, then $\mathrm{diam}(\widehat{\Phi}_H(\mu), \widehat{\Phi}_H(\mu')) < c$.

If considering $\mathcal{M}(S)$ then fix S to be either $S_{1,1}$ or $S_{0,4}$, if considering $\mathcal{P}(S)$ then fix S to be either $S_{1,2}$ or $S_{0,5}$ and use hierarchies-without-annuli.

Lemma 5.6

- When S is either $S_{1,1}$ or $S_{0,4}$, then the hierarchy paths form a coarsely contracting family of paths on $\mathcal{M}(S)$.

- When S is either $S_{1,2}$ or $S_{0,5}$, then the hierarchy-without-annuli paths form a coarsely contracting family of paths on $\mathcal{P}(S)$.

Proof There exist a hierarchy connecting any pair of points in $\mathcal{M}(S)$, which shows that the set of hierarchy paths form a transitive path family (similarly in $\mathcal{P}(S)$ for hierarchy-without-annuli paths). Fix H a hierarchy path between two points in $\mathcal{M}(S)$ and use $\widehat{\Phi}$ to denote $\widehat{\Phi}_H$.

We need to show: there exist constants b and c so that if $\mu, \mu' \in \mathcal{M}$ satisfy $d_{\mathcal{M}}(\mu, \mu') < b \cdot d_{\mathcal{M}}(\mu, \widehat{\Phi}(\mu))$, then $\mathrm{diam}(\widehat{\Phi}(\mu), \widehat{\Phi}(\mu')) < c$.

First notice that a key property of $\widehat{\Phi}$ is that there exists a constant $Q = 3M + 3\delta'$ so that every domain A of a geodesic in H (including the main surface S) has $\mathrm{diam}_A(\widehat{\Phi}(\mu) \cup p_A(\mu)) < Q$; this fact is a step in the proof of Lemma 5.4.

Fix a hierarchy G from $\widehat{\Phi}(\mu)$ to μ. Let Z be a domain of G of length at least $4Q$, an assumption which implies that:

$$(5\text{–}2) \qquad\qquad d_Z(\mu, \widehat{\Phi}(\mu)) > 3Q + M.$$

In order to compute $d_{\mathcal{M}}(\widehat{\Phi}(\mu), \widehat{\Phi}(\mu'))$ it is useful to calculate $d_Y(\mu, \mu')$ for each large domain Y of H. Consider a subsurface Z. If Z is a large domain of H, then the definition of $\widehat{\Phi}$, Lemma 3.16, and Equation (5–1) together imply that $d_Z(\mu, \partial Y) \geq d_Z(\mu, \widehat{\Phi}(\mu)) - 5M - 3\delta'$. If Z is not a large domain of H, then $d_Z(\partial Y, \widehat{\Phi}(\mu)) < 6M + 4\delta'$; applying the triangle inequality then gives $d_Z(\mu, \partial Y) \geq d_Z(\mu, \widehat{\Phi}(\mu)) - \mathrm{diam}_Z(\widehat{\Phi}(\mu)) - d_Z(\widehat{\Phi}(\mu), \partial Y) - 2\delta'$.

Combining the two cases of the previous paragraph, for any surface Z, independent of whether or not it is a large domain in H, we have:

$$(5\text{–}3) \qquad\qquad d_Z(\mu, \partial Y) \geq d_Z(\mu, \widehat{\Phi}(\mu)) - 3Q.$$

Combining the above with inequality (5–2) we have that $d_Z(\mu, \partial Y) > 3Q + M - 3Q = M$, which by the Projection estimates Theorem implies that $d_Y(\mu, \partial Z) < M$.

Now consider $d_Z(\mu', \partial Y)$: either $d_Z(\mu', \partial Y) > M$ for every large domain Y of H or for some domain Y of H we have $d_Z(\mu', \partial Y) < M$.

In the first case, from Theorem 4.3 (Projection estimates) we have $d_Y(\mu', \partial Z) < M$, which combines with the paragraph above to give $d_Y(\mu', \mu) < 2M$. At which point it is easy to check that there is a uniform bound on $d_{\mathcal{M}}(\widehat{\Phi}(\mu), \widehat{\Phi}(\mu'))$—the bound comes since $d_Y(\mu', \mu) < 2M$ implies that the pairs $(\mathcal{L}(\mu), \mathcal{R}(\mu))$ and $(\mathcal{L}(\mu'), \mathcal{R}(\mu'))$ agree (except for a switch of at most one surface which then must have diameter not much larger than the threshold for $\mathcal{G}$) and that $\Lambda(\mu)$ is coarsely the same as $\Lambda(\mu')$. Carrying out this computation, one obtains $d_{\mathcal{M}(S)}(\widehat{\Phi}(\mu), \widehat{\Phi}(\mu')) < 10M + 6\delta'$.

In the other case, when there is some domain Y of H for which $d_Z(\mu', \partial Y) < M$, we can combine this with equation (5–3) to give:

$$(5\text{–}4) \qquad\qquad d_Z(\mu, \mu') \geq d_Z(\mu, \widehat{\Phi}(\mu)) - 3Q - M.$$

This tells us that whenever $d_Z(\widehat{\Phi}(\mu), \mu) > 2(3Q + M)$, we have

$$(5\text{–}5) \qquad\qquad d_Z(\mu, \mu') > \frac{1}{2} d_Z(\widehat{\Phi}(\mu), \mu)$$

Thus, in domains Z of G for which $d_Z(\widehat{\Phi}(\mu), \mu)$ is sufficiently large, either: $d_Z(\mu', \partial Y) > M$, and thus $\widehat{\Phi}(\mu)$ and $\widehat{\Phi}(\mu')$ are close in the markings complex, or $d_Z(\mu, \mu')$ is a definite fraction of $d_Z(\widehat{\Phi}(\mu), \mu)$.

If we can show that a definite proportion of the distance travelled in $\mathcal{M}$ between $\widehat{\Phi}(\mu)$ and μ, say $\frac{d_{\mathcal{M}}(\mu, \widehat{\Phi}(\mu))}{\alpha}$, takes place in domains larger than a threshold $6Q + 2M$ then we would be done by choosing the constant b so that $b < \frac{d_{\mathcal{M}}(\mu, \widehat{\Phi}(\mu))}{\alpha}$. This would be enough, as then μ' has no chance to move far enough away from μ to have $d_Z(\mu', \partial Y) < M$, where Z is a large domain in G and Y is a large domain in H.

That there exists a uniform constant $\alpha > 0$ for which the distance $\frac{d_{\mathcal{M}}(\mu, \widehat{\Phi}(\mu))}{\alpha}$ occurs in large surfaces follows from the following counting argument. The fact that the only domains of S which support geodesics in H are the surface S and component domains of g_H has as a consequence that if a hierarchy between I and T contains geodesics in $N + 1$ different component domains, then $d_{\mathcal{C}(S)}(I, T) \geq N$. From this we see that either $\frac{1}{2} d_{\mathcal{M}(S)}(\widehat{\Phi}(\mu), \mu)$ occurs in large domains, or the sum of the lengths of geodesics shorter than $6Q + 2M$ is larger than $\frac{1}{2} d_{\mathcal{M}(S)}(\widehat{\Phi}(\mu), \mu)$. In the latter case, we then have $|g_H| > \frac{d_{\mathcal{M}(S)}(\widehat{\Phi}(\mu), \mu)}{2(6Q + 2M)}$. Thus we always have that at least $\frac{1}{2(6Q + 2M)} d_{\mathcal{M}(S)}(\widehat{\Phi}(\mu), \mu)$ occurs in the large surfaces, as claimed.

Thus, choosing $b = \frac{1}{2(6Q + 2M)}$ and $c > 10M + 6\delta'$, we have proved the theorem. $\square$

5.3 Hyperbolicity

In this section we prove Theorems 5.2 and 5.1 using the contracting properties of hierarchy paths provided by Lemma 5.6.

In [42] a contraction property is defined similar to Definition 5.5 and is used to prove δ–hyperbolicity of $\mathcal{C}(S)$; that property motivated the definition of the previous section. In this section, using results about $\mathbb{R}$–trees we give a new proof that such contraction properties imply hyperbolicity, using techniques different than those of [42]. This proof provides a new perspective concerning the relationship between stability of quasigeodesics and hyperbolicity.

Theorem 5.7 *If X is a geodesic space with a family of (K, C)–quasi-geodesic paths $\mathcal{H}$ which have the coarsely contracting property, then X is δ–hyperbolic.*

Theorems 5.2 and 5.1 follow from this theorem since Lemma 5.6 proves that in the low complexity cases $\mathcal{M}(S)$ (and $\mathcal{P}(S)$) the hierarchy paths form a coarsely contracting family of paths (respectively hierarchy-without-annuli paths), and [41, Efficiency of hierarchies Theorem] implies that hierarchy paths are (K, C)–quasi-geodesics, with the constants depending only on the topological type of the surface.

Proof   The method of proof is to show that in the asymptotic cone $\widehat{\Phi}$ induces a locally constant map to a certain family of bi-Lipschitz paths, we then show that this implies that $\mathrm{Cone}_\omega(X, \langle d_i \rangle)$ is an $\mathbb{R}$–tree. In particular we will show for each pair of points $x, y \in \mathrm{Cone}_\omega(X, \langle d_i \rangle)$ there exists a map $\Phi \colon \mathrm{Cone}_\omega(X, \langle d_i \rangle) \to [x, y]$ such that:

(1) $\Phi|_{[x,y]}$ is the identity

(2) $\Phi|_{\mathrm{Cone}_\omega(X,\langle d_i \rangle) \setminus [x,y]}$ is locally constant.

By our hypothesis, for each $H \in \mathcal{H}$ there exists a map $\widehat{\Phi}_H \colon X \to X$ such that:

(1)   For each $\mu \in X$ we have $\widehat{\Phi}_H(\mu) \in H$.

(2)   There exist constants $b, c > 0$ such that for any sequence of paths $H_n \in \mathcal{H}$ and points μ_n for which $r_n = d_X(\mu_n, \pi_H(\mu_n))$ grows linearly, then $\mathrm{diam}_X(\widehat{\Phi}_{H_n}(B_{b \cdot r_n}(\mu_n))) \leq c$.

Fix three points $\bar{I} = \langle I_i \rangle_{i \in \mathbb{N}}, \bar{T} = \langle T_i \rangle_{i \in \mathbb{N}}, \bar{\mu} = \langle \mu_i \rangle_{i \in \mathbb{N}} \in \mathrm{Cone}_\omega(X, \langle d_i \rangle)$. Let $\bar{H}$ denote the path connecting $\bar{I}$ to $\bar{T}$ given by a rescaled sequence of paths H_i connecting I_i to T_i and assume $\bar{\mu} \notin \bar{H}$. Since $\mathcal{H}$ has the coarsely contracting property, by definition we know that when $d_X(\mu_i, H_i)$ grows linearly in d_i (say it grows roughly as $g \cdot d_i$ for a constant g) the ball with radius $b \cdot d_X(\mu_i, H_i)$ around μ_i maps to a ball of radius at most c in H_i. Thus in the rescaled metric space, $\frac{1}{d_i} \cdot X$, we have $d_{\frac{1}{d_i} \cdot X}(\mu_i, H_i)$ is approximately g and in this metric the ball $B_{b \cdot g d_i}(\mu_i)$ maps to a set of diameter $\frac{c}{d_i}$ in H_i. Thus $\mathrm{diam}_{\mathrm{Cone}_\omega(X, \langle d_i \rangle)}(\langle \widehat{\Phi}(B_{b \cdot g}(\mu_i)) \rangle) = \lim_\omega \frac{c}{d_i} = 0$, thereby showing that the map is locally constant.

This proves that any embedded path connecting $\bar{I}$ and $\bar{T}$ must be a subset of $\bar{H}$. Since by hypothesis each $H \in \mathcal{H}$ is a (K, C)–quasi-geodesics, we have that $\bar{H}$ is the ultralimit as $i \to \infty$ of $(K, \frac{C}{d_i})$–quasi-geodesics and is thus a K–bi-Lipschitz embedded path. In particular it is homeomorphically embedded, so the only subpath connecting its endpoints is the path itself—thereby proving that there is a unique topological arc between any pair of points in $\mathrm{Cone}_\omega(X)$. It is worth remarking that since $\mathrm{Cone}_\omega(X)$ is a geodesics space, the path $\bar{H}$ which *a fortiori* need only be bi-Lipschitz embedded is actually the image of the geodesic between $\bar{I}$ and $\bar{T}$.

Mayer and Oversteegen's topological characterization of $\mathbb{R}$–trees [43] implies that in order to prove that $\mathrm{Cone}_\omega(X)$ is an $\mathbb{R}$–tree it suffices to show that it is uniquely arcwise connected and locally path connected. The first property was shown above, the second is true for any geodesic space.

Thus each $\mathrm{Cone}_\omega(X)$ is an $\mathbb{R}$–tree and thus Theorem 2.7 implies that X is δ–hyperbolic. $\qquad\square$

6 Asymptotic cones of $\mathcal{MCG}$

When $\xi(S) \geq 2$, the mapping class group of S is not hyperbolic and thus the results of the previous section clearly do not hold for arbitrary surfaces (similarly for the Teichmüller space of S when $\xi(S) \geq 3$). Nonetheless, in this section we show that negative curvature phenomena is present in certain directions. We begin by building a projection to mapping class groups of subsurfaces which we use to construct a subset which we call the strongly bounded subset, F_0, of the asymptotic cone. Theorem 6.5 is the main result of this section, it describes various topological properties of the subset F_0. The proof of Theorem 6.5 is similar in spirit to the proof of Theorems 5.1 and 5.2. The main difference, is that due to the lack of δ–hyperbolicity in the spaces being considered in this section, we have no hope of showing that there is a coarsely locally constant projection to every hierarchy path. We consider a subcollection of hierarchy paths for which we build such projections; it the ultralimit of this collection of path that yield the set F_0.

Given any subsurface $Z \subset S$, we define a projection $\pi_{\mathcal{M}(Z)} \colon \mathcal{M}(S) \to 2^{\mathcal{M}(Z)}$, which sends elements of $\mathcal{M}(S)$ to subsets of $\mathcal{M}(Z)$. Given any $\mu \in \mathcal{M}(S)$ we build a marking on Z in the following way. Choose an element $\gamma_1 \in \pi_Z(\mu)$, and then recursively choose γ_n from $\pi_{Z \setminus \cup_{i<n} \gamma_i}(\mu)$, for each $n \leq \xi(Z)$. Now take these γ_i to be the base curves of a marking on Z. For each γ_i we define its transversal $t(\gamma_i)$ to be $\pi_{\gamma_i}(\mu)$. This process yields a complete marking on Z, since there are $\xi(Z)$ base curves and each has a transversal; work of Masur and Minsky [41, Lemma 2.4] shows that any complete marking uniquely defines a finite subset of complete clean markings of diameter at most 6. Thus this construction yields a bounded diameter subset of $\mathcal{M}(Z)$, but arbitrary choices in choosing the γ_i were made along the way; we define $\pi_{\mathcal{M}(Z)}(\mu)$ to be the union of all possible markings built following this process, in the following lemma we show that this set has bounded diameter.

Lemma 6.1 *There exist a uniform bound depending only on S, which bounds the diameter of $\mathrm{diam}_{\mathcal{M}(Z)}(\pi_{\mathcal{M}(Z)}(\mu))$ for any $\mu \in \mathcal{M}(S)$ and $Z \subset S$.*

Proof Fix $\mu \in \mathcal{M}(S)$, $Z \subset S$, and $\nu, \nu' \in \pi_{\mathcal{M}(Z)}(\mu)$. It suffices to consider ν with base curves, γ_i, and transversals, $\pi_{\gamma_i}(\mu)$, as given above since [41, Lemma 2.4] implies that each such complete marking has a bounded diameter subset of $\mathcal{M}(Z)$ associated with it (similarly for ν'). We consider the projections of ν into $\mathcal{C}(Y)$ for each $Y \subseteq Z$ and show that the image under these projections are close to $\pi_Y(\mu)$. We proceed inductively, by considering the smallest i for which $\pi_Y(\gamma_i) \neq \varnothing$.

If $i = 1$, then since $d_{\mathcal{C}(Z)}(\gamma_1, \mu) \leq 1$, Lemma 3.3 (Lipschitz projection) implies that $d_{\mathcal{C}(Y)}(\gamma_1, \mu) \leq 3$ for any $Y \subset Z$. Thus we obtain $d_{\mathcal{C}(Y)}(\nu, \mu) \leq 3$.

We now assume the inductive hypothesis that for a fixed n we have $d_{\mathcal{C}(Y)}(\nu, \mu) \leq 3$ for all $Y \subset Z$ with $\pi_Y(\gamma_i) \neq \varnothing$ for some $i < n$. Our goal is to show that if n is the smallest index of a curve $\gamma_n \in \text{base}(\nu)$ with $\pi_Y(\gamma_n) \neq \varnothing$, then $d_{\mathcal{C}(Y)}(\nu, \mu) \leq 3$. Since $\gamma_i \cap Y = \varnothing$ when $i < n$, we have $d_{\mathcal{C}'(Y)}(\pi'_Y(\mu), \pi'_Y(\gamma_n)) \leq 1$. Thus by Lemma 3.3 we have $d_{\mathcal{C}(Y)}(\pi_Y(\mu), \pi_Y(\gamma_n)) \leq 3$, and hence $d_Y(\mu, \nu) \leq 3$. Thus for any surface $Y \subset Z$ with $\pi_Y(\text{base}(\nu)) \neq \varnothing$ we have $d_Y(\mu, \nu) \leq 3$.

It remains to deal with the case where Y is an annulus with base curve $\gamma \in \text{base}(\nu)$, but here it follows from the definition of $\pi_{\mathcal{M}(Z)}(\mu)$ that $\pi_Y(\mu) = \pi_Y(\nu)$, thereby proving that for any $\nu \in \pi_{\mathcal{M}(Z)}(\mu)$ and any $Y \subset Z$ we have $d_Y(\mu, \nu) \leq 3$. This implies that and any $Y \subset Z$ we have $\text{diam}_Y(\nu, \nu') \leq 9$.

Now, Theorem 3.17 shows that $d_{\mathcal{M}(Z)}(\nu, \nu') \sim \sum_{\substack{Y \subseteq Z \\ d_Y(\mu, \nu) > t}} d_Y(\nu, \nu')$, but we just showed that when the threshold is taken larger than 9 then the sum on the right side is zero. Thus if we choose a threshold $t > 9$ in Theorem 3.17, then $d_{\mathcal{M}(Z)}(\nu, \nu')$ is bounded by the additive quasi-isometry constant in the conclusion of that theorem. □

As in the case of the maps π_Z, we abbreviate $d_{\mathcal{M}(Z)}(\pi_{\mathcal{M}(Z)}(\mu_i), \pi_{\mathcal{M}(Z)}(\nu_i))$ by writing $d_{\mathcal{M}(Z)}(\mu_i, \nu_i)$.

Definition 6.2 Define a pair of points (μ, ν) in $\mathcal{M}_\omega(S) \times \mathcal{M}_\omega(S)$ to be of *bounded growth* if there exists a constant D for which:

$$\limsup_\omega \{d_{\mathcal{M}(Z)}(\mu_i, \nu_i) : Z \subsetneq S\} \leq D < \infty.$$

For our fixed basepoint $0 \in \mathcal{M}(S)$, let F_0 denote the set of ν such that the pair $(0, \nu)$ is of bounded growth. We use the term *strongly bounded* to refer to the elements of F_0.

Remark 6.3 Fixing a sequence of scaling factors d_n and an ultrafilter ω, one can consider those sequences $D(n)$ for which $\lim_\omega \frac{D(n)}{d_n} = 0$: such sequences are said to have *sublinear growth (with respect to the scaling constants d_n and ultrafilter ω)*. This definition allows one to consider, for instance, the set of points $\mu \in \mathcal{M}_\omega(S)$ for which the sequence $\{d_{\mathcal{M}(Z)}(\mu_i, 0) : Z \subsetneq S\}$ grows sublinearly, we call these points *strongly sublinear (for the choices d_n and ω)*. (A special case of this is for the scaling constants $d_n = \frac{1}{n}$, indeed this class of asymptotic cones was used in many of the original papers in the subject, eg, [24; 23; 53; 34].)

We note that the remaining results and proofs of this section hold mutatis mutandis replacing "strongly bounded" by "strongly sublinear." This is often useful, but in the sequel we use the strongly bounded hypothesis in order to have a condition which holds in all asymptotic cones simultaneously.

Example 6.4 If $\rho \in \mathcal{MCG}(S)$ is pseudo-Anosov, then the sequence $\langle \rho^n \rangle_n$ has distance from the origin, 0, growing at a linear rate. At the same time, there is a uniform bound M so that for all $n \in \mathbb{N}$ and $Y \subseteq S$ one has $d_Y(\rho^n, 0) \leq M$. This shows that $|F_0| > 1$. This is an important example, as in this case, up to some quasi-isometry constants, all the mapping class group distance is accounted for by distance in the complex of curves of the surface S.

This notion of degrees of boundedness provides an interesting stratification of $\mathcal{M}_\omega$; Theorems 6.5 and Theorem 7.1 are examples of how this stratification provides information about the mapping class group. These results are refined further in Behrstock–Minsky's proof of the Rank Conjecture [6].

In Section 5 we used projections from $\mathcal{M}(S)$ to hierarchy paths and showed that these maps had a strong contraction property. In fact, these maps were sufficiently contracting so that for each hierarchy path, in the asymptotic cone we obtained a locally constant map from $\mathcal{M}_\omega(S)$ to a bi-Lipschitz path. From this we concluded that in the low complexity cases $\mathcal{M}_\omega(S)$ is an $\mathbb{R}$–tree. This argument relied heavily on the fact that these surfaces had low topological complexity, nonetheless we extend the essential parts of that argument to produce strongly contracting maps in mapping class groups of surfaces of arbitrary complexity. In so doing, the following result shows that bi-Lipschitz flats in $\mathcal{M}_\omega$ intersects F_0 in a very simple way. This answers and extends Conjecture 6.4.6 of [3]. This contraction property also shows that given a pseudo-Anosov element ϕ and sufficiently large $n \in \mathbb{N}$, then any path in $\mathcal{M}(S) \setminus B_{|\phi^{2n}|}(1)$ has length which is at least quadratic in $|\phi^{2n}|$. This holds since the contraction property shows that linear length paths in this space move at most a uniformly bounded amount when projected to the quasi-geodesic from $\phi^{-n}(1)$ to $\phi^n(1)$, therefore any path in $\mathcal{M}(S) \setminus B_{|\phi^{2n}|}(1)$ connecting $\phi^{-2n}(1)$ to $\phi^{2n}(1)$ must contain at least linearly many disjoint linear length subpaths. Except when it is virtually free, $\mathcal{M}(S)$ is thick of order 1 [4] and thus has at most quadratic divergence. Hence, except in those cases where it is virtually free, the mapping class group has a quadratic divergence function. See the next subsection for some further topological corollaries and a discussion of other applications.

Theorem 6.5 F_0 *satisfies the following properties:*

(1) F_0 *is geodesically convex.*

(2) *Any distinct pair of points in F_0 are connected by a unique path in $\mathcal{M}_\omega$. In particular, F_0 is an $\mathbb{R}$–tree.*

(3) *In $\mathcal{M}_\omega$, any bi-Lipschitz flat of rank larger than one intersects F_0 in at most one point.*

Before proving this theorem we establish some notation and provide a lemma which we shall need in the proof.

Similar to the definition in Section 5, for a hierarchy G we define the *large domains* of G as:

$$\mathcal{G}_{G,C} = \{Y \subseteq S : d_Y(I(G), T(G)) > C \geq 3M\}.$$

We will drop the C from the notation when the choice of C is irrelevant. Also, we define the *initial large domains* by:

$$I_G = \{Y \in \mathcal{G}_G : \text{Each } Y' \in \mathcal{G} \text{ which overlaps } Y \text{ has } Y \prec_t Y'\}.$$

From finiteness of $\mathcal{G}$ it follows $I_G \neq \varnothing$; since any two domains of I_G are disjoint or nested there is a bound on $|I_G|$ depending only on the topological type of S.

Lemma 6.6 *If $Y' \in \mathcal{G}_G \setminus I_G$ then there exists $Y \in I_G$ such that Y' overlaps Y.*

Proof Letting $Y' \notin I_G$, then there must exist a domain $Y_1 \in \mathcal{G}_G$ which overlaps Y' and for which $Y_1 \prec_t Y'$. Either $Y_1 \in I_G$ in which case we are finished, or we can find a surface $Y_2 \in \mathcal{G}_G$ which overlaps Y_1 and for which $Y_2 \prec_t Y_1$. Since $Y_1 \in \mathcal{G}$, Lemma 3.16 implies that $d_{Y_1}(\partial Y', \partial Y_2) > M$ and thus Y' must overlap Y_2. Repeating this process finitely many times (since $\mathcal{G}_G$ is finite) yields a sequence $Y \prec_t \ldots \prec_t Y_2 \prec_t Y_1 \prec_t Y'$ with $Y \in I_G$ and Y overlapping Y'. $\qquad\square$

We now proceed to the proof of Theorem 6.5.

Proof Fix two points $x = \langle x_i \rangle, y = \langle y_i \rangle \in F_0$. Fix a hierarchy H_i connecting x_i to y_i and let $[x_i, y_i]$ denote the corresponding hierarchy path between them. We write $[x, y]$ to denote the bi-Lipschitz path in $\mathcal{M}_\omega(S)$ obtained by taking an ultralimit of these quasi-geodesic paths.

(1) We will now show that every point of $[x, y]$ is in F_0. By the definition of F_0, for every $\langle \alpha_i \rangle \in \text{Seq} \setminus \langle S \rangle$ we have

$$\lim_\omega (\max\{d_{\mathcal{M}(\alpha_i)}(x_i, 0), d_{\mathcal{M}(\alpha_i)}(y_i, 0)\}) < \infty.$$

Thus, in particular, for each $\langle \alpha_i \rangle \in \text{Seq} \setminus \langle S \rangle$ we have

$$(6\text{--}1) \qquad \lim_\omega \frac{1}{d_i} d_{\mathcal{M}(\alpha_i)}(x_i, y_i) = 0.$$

We note the following useful property of hierarchy paths: if $[\mu, \nu]$ is a hierarchy path from μ to ν and $\rho \in [\mu, \nu]$, then for all $Y \subsetneq S$ one has $d_{\mathcal{C}(Y)}(\nu, \rho) \leq d_{\mathcal{C}(Y)}(\nu, \mu) + 3M$ (where M is the constant of Theorem 4.3, depending only on S). Invoking Theorem 3.17 and Lemma 6.1 this implies that $d_{\mathcal{M}(Y)}(\nu, \rho) < K d_{\mathcal{M}(Y)}(\nu, \mu) + C$, for constants K and C which depend only on S.

By definition of the asymptotic cone, if $\langle w_i \rangle \in [x, y]$, then for an ω–large subsequence of indices the distances $d_{\mathcal{M}(S)}(w_i, [x_i, y_i])$ grow slower than the sequence d_n; as a consequence of the previous paragraph we see that for any $\langle w_i \rangle \in [x, y]$ and any $\langle \alpha_i \rangle \in \text{Seq} \setminus \langle S \rangle$ we have $d_{\mathcal{M}(\alpha_i)}(w_i, y_i) \leq K d_{\mathcal{M}(\alpha_i)}(x_i, y_i) + C$ for ω–almost every i. Together with equation (6–1) this implies that for any $\langle \alpha_i \rangle \in \text{Seq} \setminus \langle S \rangle$ we have $\lim_\omega \frac{1}{d_i} d_{\mathcal{M}(\alpha_i)}(y_i, w_i) = 0$. Namely, we have shown that $[x, y] \subset F_0$, ie, F_0 is geodesically convex.

(2) This argument follows the general outline used to prove Lemma 5.6. Consider the family $\mathcal{H}$ consisting of ultralimits of rescaled hierarchy paths connecting pairs of points in F_0. We shall show that for any $x, y \in F_0$ and any path $[x, y] \in \mathcal{H}$ there is a continuous map $\Phi \colon \mathcal{M}_\omega(S) \to [x, y]$ such that:

(1) $\Phi|_{[x,y]}$ is the identity

(2) Φ is locally constant outside $[x, y]$.

Once this is established Theorem 5.7, which we previously used in conjunction with Lemma 5.6 to prove Theorem 5.1, implies that F_0 is an $\mathbb{R}$–tree. The two key differences between this argument and that of Lemma 5.6 are in the way in which the map Φ is constructed and then the proof that this map is locally constant—since the present construction of Φ works for surfaces of arbitrary complexity we must rely on the property of being strongly bounded for both these steps. Similar to the construction of Φ in Section 5, we define our projection by considering a sequence of hierarchies and maps to subsets of their slices and then perform a limiting procedure.

Φ **is well-defined** Fix $\mu \in \mathcal{M}_\omega$ and let $g_{H_i} \subset \mathcal{C}(S)$ denote the main geodesic of H_i. By δ–hyperbolicity of $\mathcal{C}(S)$, *a la* Lemma 2.3, the closest point projection of $\pi_S(\mu_i)$ to g_{H_i} determines a subset $C_{\mu_i} \subset g_{H_i}$ whose diameter is uniformly bounded by a constant $K(\delta)$, which depends only on the δ–hyperbolicity constant of $\mathcal{C}(S)$. Define $\widehat{\Phi}(\mu_i)$ to be the set of all marking in $[x_i, y_i]$ whose base curves include any element $\gamma \in C_{\mu_i}$.

Note that given any element $\gamma \in C_{\mu_i}$, by strong boundedness there are at most a bounded number of elements of $\mathcal{C}(\gamma)$ which appear as transverse curves to γ in some marking of H_i. Further, strong boundedness also implies there is a uniform bound on $\mathrm{diam}_{\mathcal{C}(S \setminus \gamma)}([x_i, y_i])$. Thus, inductively building the subset $\Phi(\mu) \subset [x_i, y_i]$ one sees that there is a uniform bound, K, on its diameter.

In particular, since for each $\mu_i \in [x_i, y_i]$ one has $\mu_i \in \Phi(\mu_i)$, the above uniform bound implies that $\mathrm{diam}_{\mathcal{M}_\omega}(\langle \Phi(\mu_i) \rangle, \langle \mu_i \rangle) = \lim_\omega \frac{1}{d_i}(K) = 0$. Thus one obtains that $\Phi(\mu)$ is always a point, this also shows that Φ restricts to the identity map on $[x, y]$.

Φ is locally constant off $[x, y]$ Fix a point $\mu \in \mathcal{M}_\omega(S) \setminus [x, y]$ and a point $\mu' \in \mathcal{M}_\omega(S)$ with $d(\mu, \mu') < c \cdot d(\mu, \Phi(\mu))$ for a positive constant c, depending only on the topological type of S, to be determined by the following argument. We want to show that for any such pair one obtains $\Phi(\mu) = \Phi(\mu')$.

Fix a representative for $\mu = \langle \mu_i \rangle$ and for each i consider a hierarchy G_i whose initial marking is a point of $\widehat{\Phi}(\mu_i)$ and with terminal marking μ_i—this involves an arbitrary choice of element of $\widehat{\Phi}(\mu_i)$, but the fact that $\mathrm{diam}(\widehat{\Phi}(\mu_i))$ is uniformly bounded implies that any choices here yield the same point in the asymptotic cone. As usual, g_{G_i} will denote the main geodesic of G_i.

The remained of the argument will be divided into two cases depending on the length of g_{G_i}. Let K denote a fixed constant coming from Theorem 3.17 (K is chosen by taking the quasi-isometry constant given by Theorem 3.17 corresponding to a threshold $t > 4M + D$, where M is the constant given by Theorem 4.3 and D is the constant quantifying the strong boundedness of the pair (x, y), as given in Definition 6.2).

Case (i) $|g_G| > \frac{1}{K} d(\mu, \Phi(\mu))$ More precisely, here we assume that for ω–almost every i the length of g_{G_i} is larger than $\frac{1}{K} d_{\mathcal{M}(S)}(\mu_i, \widehat{\Phi}(\mu_i))$.

Let $c < \frac{1}{2K}$. Hyperbolicity of $\mathcal{C}(S)$, then implies that there is a uniform bound on $\mathrm{diam}_{\mathcal{C}(S)}(C(\mu_i) \cup C(\mu_i'))$ (see Theorem 3.2 and Lemma 2.3). The proof that Φ is well defined, then implies that $\Phi(\mu) = \Phi(\mu')$.

Case (ii) $|g_G| \leq \frac{1}{K} d(\mu, \Phi(\mu))$ We now consider the cases not covered by Case (i), namely we assume that for ω–almost every i the length of g_{G_i} is at most $\frac{1}{K} d_{\mathcal{M}(S)}(\mu_i, \widehat{\Phi}(\mu_i))$. The choice of K implies that for ω–almost every i, any hierarchy between μ_i and $\widehat{\Phi}(\mu_i)$ must have a domain $Y \subsetneq S$ for which $d_Y(\mu_i, \widehat{\Phi}(\mu_i)) > 4M + D$, where M is the constant given by Theorem 4.3 and D is the strongly boundedness constant for the pair (x, y). Fix a constant $0 < c < 1$ and let $\mu_i =$

$\mu_{i,0}, \mu_{i,1}, \mu_{i,2}, \ldots, \mu_{i,k} = \mu_i'$ be an elementary move sequence of a hierarchy path in $\mathcal{M}(S)$ with $k < cd_{\mathcal{M}(S)}(\mu_i, \widehat{\Phi}(\mu_i))$.

For any surface $Y \subsetneq S$, define

$$J_{i,Y} = \{j \in [0, k] : d_Y(\mu_{i,j}, \mu_i) > 4M + D \text{ and } d_Y(\mu_{i,j}, \widehat{\Phi}(\mu_i)) > 4M + D\}.$$

We now give a brief digression to study properties of $J_{i,Y}$. Suppose $j \in J_{i,Y}$, $l \in J_{i,Z}$, and $Y, Z \subsetneq S$ overlap. Then, Y and Z are both large domains for G_i, and in the time ordering on the hierarchy G_i either $Y \prec_t Z$ or $Z \prec_t Y$, for the remainder of the paragraph we assume the first ordering (in the other case an analogous argument can easily be provided by the reader). Lemma 3.16 shows that $Y \prec_t Z$ implies $d_Y(\partial Z, \mu_i) < M$ and $d_Z(\partial Y, \widehat{\Phi}(\mu_i)) < M$; the first inequality combines with the definition of $J_{i,Y}$ to imply that $d_Y(\mu_{i,j}, \partial Z) > 3M + D$. Theorem 4.3 then implies that $d_Z(\mu_{i,j}, \partial Y) < M$, which combines with the second inequality of the previous sentence to show that $d_Z(\mu_{i,j}, \widehat{\Phi}(\mu_i)) < 2M$. Thus $j \notin J_{i,Z}$ and we have shown that for any $Y, Z \subsetneq S$ we have $J_{i,Y} \cap J_{i,Z} = \varnothing$. Thus $J_{i,Y}$ and $J_{i,Z}$ share a point only if they are disjoint or nested, a straightforward topological count gives a bound of $2\xi(S)$ for the number of $J_{i,Y}$ which can cover a given j. Furthermore, this argument shows that in any hierarchy with initial marking μ_i, then either $d_{\mathcal{C}(Y)}(\mu_i, \mu_{i,j}) < 2M$ for all j or the projection to $\mathcal{C}(Z)$ of the terminal marking of this hierarchy is within M of the point $\widehat{\Phi}(\mu_i)$; when the later condition is satisfied we say that this hierarchy has *traversed* Z.

What we shall now show is that the hierarchy from μ_i to μ_i' can not possibly traverse all the large domains appearing in the hierarchy G_i. This then implies that there exists a large domain of G_i, which we will call Y_i', in which $d_{\mathcal{C}(Y_i')}(\mu, \mu_i') < 2M$. The strongly bounded hypothesis implies that $\text{diam}_{\mathcal{C}(Y_i')}([x_i, y_i]) < K$, together with the fact that $d_{\mathcal{C}(Y_i')}(\mu, \widehat{\Phi}(\mu_i)) > 4M + D$ this allows us to conclude that the main geodesic of any hierarchy from μ_i' to $[x_i, y_i]$ must pass through the vertex $\partial Y_i' \in \mathcal{C}(S)$. From this and the fact that $[x_i, y_i] \subset F_0$ we then give a quick deduction of the theorem.

Recall that $\mathcal{G}_{G_i, 4M+D}$ denotes the set of domains in G_i supporting geodesics larger than $4M + D$; unless noted otherwise all time orderings on domains will refer to the ordering in the hierarchy G_i. Recall that the set of large initial domains for G_i has cardinality at most $2\xi(S)$. Thus, by Lemma 6.6 there exists $Y_i' \subsetneq S$ with $Y_i' \in \mathcal{G}_{G_i, 4M+D}$, such that if we define

$$\mathcal{I}_i = \{Y_i'\} \cup \{Z \in \mathcal{G}_{G_i, 4M+D} : Y_i' \prec_t Z\},$$

then for ω–almost every i we have $\displaystyle\sum_{Z\in\mathcal{I}_i} d_{\mathcal{C}(Z)}(\mu_i, \widehat{\Phi}(\mu_i)) > \alpha d_{\mathcal{M}(S)}(\mu_i, \widehat{\Phi}(\mu_i))$,

for a fixed $\alpha > 0$ independent of i.

Since the $J_{i,Z}$ cover the path from μ_i to μ_i' with degree at most $2\xi(S)$, we have $2k\xi(S) \geq \sum_{Z\in\mathcal{I}_i} |J_{i,Z}|$. Now we use the result from [41, page 962] that there exists a constant $\beta > 0$ (depending only on the topological type of S) for which $|J_{i,Z}| \geq \beta d_{\mathcal{C}(Z)}(\mu_i, \widehat{\Phi}(\mu_i))$ if the path from μ_i to μ_i' traverses the large domain Z. Thus, if all the domains $Z \in \mathcal{I}_i$ are traversed between μ_i and μ_i', we have

$$2k\xi(S) \geq \sum_{Z\in\mathcal{I}_i} |J_{i,Z}| \geq \sum_{Z\in\mathcal{I}_i} \beta d_{\mathcal{C}(Z)}(\mu_i, \widehat{\Phi}(\mu_i)) \geq \alpha\beta d_{\mathcal{M}(S)}(\mu_i, \widehat{\Phi}(\mu_i)).$$

Thus, $k \geq \frac{\alpha\beta d_{\mathcal{M}(S)}(\mu_i, \widehat{\Phi}(\mu_i))}{2\xi(S)}$; taking $c \leq \frac{\alpha\beta}{2\xi(S)}$ we obtain a contradiction, and hence we may assume that not all the large domains in $\mathcal{I}_i$ are traversed between μ_i and μ_i'.

In particular, we may assume Y_i' can not be completely traversed between μ_i and μ_i' and is thus a large domain for any hierarchy from μ_i' to $\widehat{\Phi}(\mu_i)$. Therefore, the main geodesic of any hierarchy from μ_i' to $\widehat{\Phi}(\mu_i)$ contains the vertex $\partial Y_i'$. If $d_{\mathcal{C}(S)}(\partial Y_i', \widehat{\Phi}(\mu_i)) > \delta$, then any $\mathcal{C}(S)$ geodesic from $\pi_{\mathcal{C}(S)}(\mu_i')$ to the main geodesic of H_i must pass within δ of $\partial Y_i'$. This yields a uniform bound on $\mathrm{diam}(C_{\mu_i} \cup C_{\mu_i'})$ and the same argument given before shows that in this case $\Phi(\mu) = \Phi(\mu')$.

Thus, we may assume that the only choices for Y_i' are those which satisfy $d_{\mathcal{C}(S)}(\partial Y_i', \widehat{\Phi}(\mu_i)) \leq \delta$. It then follows from the Move distance and projections Theorem that for ω–almost every i there exists a domain Q_i for which $d_{\mathcal{C}(S)}(\partial Q_i, \widehat{\Phi}(\mu_i)) \leq \delta$ and with $\lim_\omega d_{\mathcal{C}(Q_i)}(\mu_i, \widehat{\Phi}(\mu_i)) \to \infty$. On the other hand, the strong boundedness of the pair (x, y) implies that $\mathrm{diam}_{\mathcal{C}(Q_i)}([x_i, y_i]) < D$ for ω–almost every i, hence by Lemma 3.14, any hierarchy between μ_i' and any point on $[x_i, y_i]$ must have a geodesic supported on Q_i. Thus the main geodesic of any hierarchy between μ_i' and any point on $[x_i, y_i]$ must pass through the point ∂Q_i. This provides a uniform bound on the distance between $\widehat{\Phi}(\mu_i)$ and $\widehat{\Phi}(\mu_i')$ and therefore shows that $\Phi(\mu) = \Phi(\mu')$, completing the proof.

(3) Any bi-Lipschitz flat which intersects F_0 in more than one point contains an arc of F_0. If this flat has rank greater than one, then this arc must be part of an embedded circle. But this would contradict the fact that there is a unique path in $\mathcal{M}_\omega$ between each pair of points of F_0.

Thus a bi-Lipschitz flat of rank greater than one can intersect F_0 in at most one point. $\square$

7 Cone$_\omega(\mathcal{MCG})$ is tree-graded

We begin by stating a consequence of Theorem 6.5 and use this as a segue to mention some interesting questions about the mapping class group. In Theorem 6.5 we obtain a unique path in $\mathcal{M}_\omega$ between any pair of points in F_0. From this it follows that the removal of any point of F_0 disconnects $\mathcal{M}_\omega$, ie, if S is any surface and $\mathcal{M}_\omega(S)$ is any asymptotic cone of $\mathcal{MCG}(S)$, then $\mathcal{M}_\omega(S)$ has a global cut-point. Since asymptotic cones of groups are homogeneous, this can be stated as:

Theorem 7.1 *If S is any surface and $\mathcal{M}_\omega(S)$ is any asymptotic cone of $\mathcal{MCG}(S)$, then every point of $\mathcal{M}_\omega(S)$ is a global cut-point.*

We now mention the following definition which was introduced in [19].

Definition 7.2 Let $\mathbb{F}$ be a complete geodesic metric space and let $\mathcal{P}$ be a collection of closed geodesic subsets (called *pieces*). The space $\mathbb{F}$ is said to be *tree-graded with respect to $\mathcal{P}$* when the following two properties are satisfied:

- The intersection of each pair of pieces has at most one point.
- Every simple geodesic triangle in $\mathbb{F}$ is contained in one piece.

In Remark 2.31 of [19], it is noted "every asymptotic cone of a group which has a cut-point is tree graded with respect to a uniquely determined collection of pieces each of which is either a singleton or a closed geodesic subset without cut-points." In light of this remark, Theorem 7.1 implies:

Corollary 7.3 *For every surface S, each asymptotic cone, $\mathcal{M}_\omega(S)$, is tree-graded.*

This result raises the question of describing the pieces. A useful object in the study of tree-graded structures is the *transversal tree* containing a point μ: this is defined to be the maximal tree (possibly consisting of only the point μ), satisfying the property that it intersects each piece in at most one point. One consequence of Theorem 6.5 is that F_0 is a subset of the transversal tree containing the basepoint 0. (Similarly by defining F_μ for any $\mu \in \mathcal{M}_\omega(S)$, we obtain a subset of the transversal tree containing the point μ.)

One of the questions solved by [19] is a proof of quasi-isometric rigidity for groups whose asymptotic cones admit tree-gradings of a certain form, namely the asymptotic cones of their peripheral subgroups are cut-point free (see [4] for a generalization). This approach might shed light on the following, which remains a major outstanding conjecture in the study of the mapping class group:

Conjecture 7.4 *Let S be a non-exceptional surface of finite type. For any finitely generated group G quasi-isometric to the mapping class group $\mathcal{MCG}(S)$, there exists a homomorphism $G \to \mathcal{MCG}(S)$ with finite kernel and finite index image.*

Informally, this question asks if the mapping class group is uniquely determined among finitely generated groups by its large scale geometry. (See [46] for a further discussion and a proof of Conjecture 7.4 for mapping class groups of once punctured surfaces.)

Another interesting question is to determine: to what extent are mapping class groups relatively hyperbolic. There are several competing definitions for relative hyperbolicity, one of which is that the Cayley graph "electrified" over the cosets of a finite collection of subgroups is δ–hyperbolic, this is now often called *weak relative hyperbolicity*. Another is weak relative hyperbolicity combined with a hypothesis called the Bounded Coset Penetration Property (BCP), which places restrictions on how paths can travel between these cosets; this condition was first formulated by Farb [20] and is often called *(strong) relative hyperbolicity*. Many people have since studied another equivalent versions of relative hyperbolicity which was formulated by Gromov [22] and then elaborated on by Bowditch, Osin, Szczepański, and others [10; 48; 51].

Another main result in [19] is that a group, G, is relatively hyperbolic with respect to a collection of subgroups, $H_1, \ldots, H_n$, which satisfy BCP if and only if $\mathrm{Cone}_\omega(G)$ is tree-graded with respect to ultralimits of cosets of the subgroups H_i. They ask the open question of whether any group whose asymptotic cones have cut-points is relatively hyperbolic (See Problem 1.18 of [19]). Masur and Minsky showed that $\mathcal{MCG}(S)$ is weakly hyperbolic relative to stabilizers of multicurves [42], thus one knows that the maximal subsets without cut-points in $\mathcal{M}_\omega(S)$ are not ultralimits of these subgroups. It is easy to verify that multicurve stabilizers do not satisfy BCP, nonetheless, there could be another collection of subgroups with respect to which the mapping class group is relatively hyperbolic. Accordingly, in an earlier version of this paper we asked the fundamental question:

Question 7.5 *Is there a collection of subgroups of $\mathcal{MCG}(S)$ with respect to which this group is strongly relatively hyperbolic?*

An answer to this question can now be given. The first author in joint work with Druţu and Mosher [4] and independently work of Anderson, Aramayona, and Shackleton [1] gives a complete answer to this question by showing that for no family of subgroups is the mapping class group strongly relatively hyperbolic.

Corollary 7.3 contrasts with the result of [4] and [1] to yield an example of a group whose asymptotic cone is tree-graded, but the group is not strongly relatively hyperbolic

for any choice of subgroups; this answers Problem 1.18 of [19] and does so with a finitely presented group. In fact, in [4] a stronger result is shown, namely, there is no collection of subsets of the mapping class group for which $\text{Cone}_\omega(\mathcal{MCG})$ is tree-graded with pieces coming from ultralimits of these subsets; thus the result that $\text{Cone}_\omega(\mathcal{MCG})$ is, nonetheless, tree-graded answers another question of [19] concerning whether pieces in the asymptotic cone always arise as ultralimits of subsets.

We note the following question which this raises.

Question 7.6 *What other groups besides $\mathcal{MCG}(S)$ have tree-graded asymptotic cones, but are not strongly relatively hyperbolic with respect to any subgroups? . . . subsets?*

After learning of our methods, B. Kleiner informed us that using analogous methods one can show that graph manifolds satisfy the property of having tree-graded asymptotic cones, but are not strongly relatively hyperbolic with respect to any subgroups [37]. (See [35] for a related discussion. See also [4] for a proof of the latter fact.)

The results about asymptotic cones throughout this work have all been independent of the choice of ultrafilter. In [39] it is shown that if one assumes the negation of the Continuum Hypothesis, then there are uniform lattices in certain semi-simple Lie groups with 2^{2^ω} different asymptotic cones up to homeomorphism (here ω denotes the first uncountable ordinal). Even in the presence of the Continuum Hypothesis, [19] produce examples of finitely generated groups with 2^ω non-homeomorphic asymptotic cones. These results make natural the following two questions.

Question 7.7

   (1) *How many non-homeomorphic asymptotic cones does the mapping class group have?*

   (2) *(Sapir) How many non-bi-Lipschitz equivalent asymptotic cones does the mapping class group have?*

Note that for the examples given in [19] of finitely generated groups with 2^ω non-homeomorphic asymptotic cones, the fundamental group can be used to distinguish the different asymptotic cones. Producing non-homeomorphic asymptotic cones for the mapping class group would require more subtle techniques since automaticity of the mapping class group implies that every asymptotic cone for this group is simply connected—this follows from results in [49] and [47].

References

[1] **J Anderson**, **J Aramayona**, **K Shackleton**, *A simple criterion for non-relative hyperbolicity and one-endedness of groups* `arXiv:math.GT/0504271`

[2] **J Aramayona**, *The Weil-Petersson geometry of the five-times punctured sphere*, PhD thesis, University of Southampton (2004))

[3] **J Behrstock**, *Asymptotic geometry of the mapping class group and Teichmüller space*, PhD thesis, SUNY at Stony Brook (2004)) Available at `http://www.math.columbia.edu/~jason`

[4] **J Behrstock**, **C Druţu**, **L Mosher**, *Thick metric spaces, relative hyperbolicity, and quasi-isometric rigidity* `arXiv:math.GT/0512592`

[5] **J Behrstock**, **D Margalit**, *Curve complexes and finite index subgroups of mapping class groups*, Geom. Dedicata 118 (2006) 71–85 MR2239449

[6] **J Behrstock**, **Y N Minsky**, *Dimension and rank for mapping class groups* `arXiv: math.GT/0512352`

[7] **M Bestvina**, **K Fujiwara**, *Bounded cohomology of subgroups of mapping class groups*, Geom. Topol. 6 (2002) 69–89 MR1914565

[8] **J S Birman**, **A Lubotzky**, **J McCarthy**, *Abelian and solvable subgroups of the mapping class groups*, Duke Math. J. 50 (1983) 1107–1120 MR726319

[9] **B H Bowditch**, *Intersection numbers and the hyperbolicity of the curve complex* Available at `http://www.maths.soton.ac.uk/staff/Bowditch/preprints.html`

[10] **B H Bowditch**, *Relatively hyperbolic groups* Available at `http://www.maths.soton.ac.uk/staff/Bowditch/preprints.html`

[11] **B H Bowditch**, *Tight geodesics in the curve complex* Available at `http://www.maths.soton.ac.uk/staff/Bowditch/preprints.html`

[12] **M R Bridson**, *Asymptotic cones and polynomial isoperimetric inequalities*, Topology 38 (1999) 543–554 MR1670404

[13] **M R Bridson**, **A Haefliger**, *Metric spaces of non-positive curvature*, Grundlehren series 319, Springer, Berlin (1999) MR1744486

[14] **J F Brock**, *The Weil-Petersson metric and volumes of 3-dimensional hyperbolic convex cores*, J. Amer. Math. Soc. 16 (2003) 495–535 MR1969203

[15] **J F Brock**, **R Canary**, **Y N Minsky**, *The classification of Kleinian surface groups II: The ending lamination conjecture* `arXiv:math.GT/0412006`

[16] **J F Brock**, **B Farb**, *Curvature and rank of Teichmüller space*, Amer. J. Math. 128 (2006) 1–22

[17] **J Burillo**, *Dimension and fundamental groups of asymptotic cones*, J. London Math. Soc. (2) 59 (1999) 557–572 MR1709665

[18] **C Druţu**, *Quasi-isometry invariants and asymptotic cones*, Internat. J. Algebra Comput. 12 (2002) 99–135   MR1902363   International Conference on Geometric and Combinatorial Methods in Group Theory and Semigroup Theory (Lincoln, NE, 2000)

[19] **C Druţu**, **M Sapir**, *Tree-graded spaces and asymptotic cones of groups*, Topology 44 (2005) 959–1058   MR2153979With an appendix by Denis Osin and Sapir

[20] **B Farb**, *Relatively hyperbolic groups*, Geom. Funct. Anal. 8 (1998) 810–840   MR1650094

[21] **É Ghys**, **P de la Harpe**, *Infinite groups as geometric objects (after Gromov)*, from: "Ergodic theory, symbolic dynamics, and hyperbolic spaces (Trieste, 1989)", Oxford Sci. Publ., Oxford Univ. Press, New York (1991) 299–314   MR1130180

[22] **M Gromov**, *Hyperbolic groups*, from: "Essays in group theory", Math. Sci. Res. Inst. Publ. 8, Springer, New York (1987) 75–263   MR919829

[23] **M Gromov**, *Asymptotic invariants of infinite groups*, from: "Geometric group theory, Vol. 2 (Sussex, 1991)", London Math. Soc. Lecture Note Ser. 182, Cambridge Univ. Press, Cambridge (1993) 1–295   MR1253544

[24] **M Gromov**, *Groups of polynomial growth and expanding maps*, Inst. Hautes Études Sci. Publ. Math. (1981) 53–73   MR623534

[25] **U Hamenstädt**, *Geometry of the mapping class groups III: Geometric rank*   arXiv: math.GT/0512429

[26] **J L Harer**, *The virtual cohomological dimension of the mapping class group of an orientable surface*, Invent. Math. 84 (1986) 157–176   MR830043

[27] **W J Harvey**, *Boundary structure of the modular group*, from: "Riemann surfaces and related topics: Proceedings of the 1978 Stony Brook Conference (State Univ. New York, Stony Brook, N.Y., 1978)", Ann. of Math. Stud. 97, Princeton Univ. Press, Princeton, N.J. (1981) 245–251   MR624817

[28] **A Hatcher**, *Pants decompositions of surfaces*   Available   at   http://www.math.cornell.edu/~hatcher

[29] **J Hempel**, *3-manifolds as viewed from the curve complex*, Topology 40 (2001) 631–657   MR1838999

[30] **Z Huang**, *Asymptotic flatness of the Weil-Petersson metric on Teichmüller space*, Geom. Dedicata 110 (2005) 81–102   MR2136021

[31] **Y Imayoshi**, **M Taniguchi**, *An introduction to Teichmüller spaces*, Springer, Tokyo (1992)   MR1215481   Translated and revised from the Japanese by the authors

[32] **E Irmak**, *Superinjective simplicial maps of complexes of curves and injective homomorphisms of subgroups of mapping class groups*, Topology 43 (2004) 513–541   MR2041629

[33] **N V Ivanov**, *Mapping class groups*, from: "Handbook of geometric topology", North-Holland, Amsterdam (2002) 523–633   MR1886678

[34] **M Kapovich**, *Lectures on the Geometric Group Theory* (2003) Available at http://www.math.ucdavis.edu/~kapovich/eprints.html

[35] **M Kapovich**, **B Kleiner**, **B Leeb**, *Quasi-isometries and the de Rham decomposition*, Topology 37 (1998) 1193–1211 MR1632904

[36] **M Kapovich**, **B Leeb**, *Quasi-isometries preserve the geometric decomposition of Haken manifolds*, Invent. Math. 128 (1997) 393–416 MR1440310

[37] **B Kleiner**, *personal communcation*

[38] **B Kleiner**, **B Leeb**, *Rigidity of quasi-isometries for symmetric spaces and Euclidean buildings*, Inst. Hautes Études Sci. Publ. Math. (1997) 115–197 (1998) MR1608566

[39] **L Kramer**, **S Shelah**, **K Tent**, **S Thomas**, *Asymptotic cones of finitely presented groups*, Adv. Math. 193 (2005) 142–173 MR2132762

[40] **D Margalit**, *Automorphisms of the pants complex*, Duke Math. J. 121 (2004) 457–479 MR2040283

[41] **H A Masur**, **Y N Minsky**, *Geometry of the complex of curves. II. Hierarchical structure*, Geom. Funct. Anal. 10 (2000) 902–974 MR1791145

[42] **H A Masur**, **Y N Minsky**, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math. 138 (1999) 103–149 MR1714338

[43] **J C Mayer**, **L G Oversteegen**, *A topological characterization of $\mathbf{R}$-trees*, Trans. Amer. Math. Soc. 320 (1990) 395–415 MR961626

[44] **Y N Minsky**, *The classification of Kleinian surface groups I: models and bounds* arXiv:math.GT/0302208

[45] **Y N Minsky**, *Quasi-projections in Teichmüller space*, J. Reine Angew. Math. 473 (1996) 121–136 MR1390685

[46] **L Mosher**, *Homology and dynamics in quasi-isometric rigidity of once-punctured mapping class groups* arXiv:math.GR/0308065 Lecture Notes from "LMS Durham Symposium: Geometry and Cohomology in Group Theory, University of Durham, UK (July 2003)

[47] **L Mosher**, *Mapping class groups are automatic*, Ann. of Math. (2) 142 (1995) 303–384 MR1343324

[48] **D V Osin**, *Relatively hyperbolic groups: intrinsic geometry, algebraic properties, and algorithmic problems*, Mem. Amer. Math. Soc. 179 (2006) vi+100 MR2182268

[49] **P Papasoglu**, *On the asymptotic cone of groups satisfying a quadratic isoperimetric inequality*, J. Differential Geom. 44 (1996) 789–806 MR1438192

[50] **T R Riley**, *Higher connectedness of asymptotic cones*, Topology 42 (2003) 1289–1352 MR1981358

[51] **A Szczepański**, *Relatively hyperbolic groups*, Michigan Math. J. 45 (1998) 611–618 MR1653287

[52]   **W P Thurston**, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bull. Amer. Math. Soc. (N.S.) 19 (1988) 417–431   MR956596

[53]   **L van den Dries**, **A J Wilkie**, *Gromov's theorem on groups of polynomial growth and elementary logic*, J. Algebra 89 (1984) 349–374   MR751150

[54]   **S A Wolpert**, *Geometry of the Weil-Petersson completion of Teichmüller space*, from: "Surveys in differential geometry, Vol. VIII (Boston, 2002)", Surv. Differ. Geom., VIII, Int. Press, Somerville, MA (2003) 357–393   MR2039996

Department of Mathematics, Barnard College
Columbia University, New York, NY 10027, USA

jason@math.columbia.edu