A N N A L E S
HENRI LEBESGUE

MLADEN BESTVINA

KEN BROMBERG

KOJI FUJIWARA

PROPER ACTIONS ON FINITE PRODUCTS OF QUASI-TREES

ACTIONS PROPRES SUR DES PRODUITS FINIS DE QUASI-ARBRES

ABSTRACT. — We say that a finitely generated group G has property (QT) if it acts isometrically on a finite product of quasi-trees so that orbit maps are quasi-isometric embeddings. A quasi-tree is a connected graph with path metric quasi-isometric to a tree, and product spaces are equipped with the ℓ^1-metric.

We prove that residually finite hyperbolic groups and mapping class groups have (QT).

RÉSUMÉ. — Un groupe de type fini G a la propriété (QT) s'il agit isométriquement sur un produit fini de quasi-arbres de telle sorte que les applications orbitales soient des plongement quasi-isométriques. Un quasi-arbre est un graphe connexe muni de la distance des chemins qui est quasi-isométrique à un arbre, et les espaces produits sont munis de la distance ℓ^1.

Nous montrons que les groupes hyperboliques résiduellement finis et les groupes modulaires ont (QT).

1. Introduction

We say that a finitely generated group G has property (QT) if it acts isometrically on a finite product of quasi-trees so that orbit maps are quasi-isometric embeddings. A quasi-tree is a connected graph with path metric quasi-isometric to a tree, and product spaces are equipped with the ℓ^1-metric. The first examples of such groups come with proper actions on products of trees, for example free groups, surface groups (e.g. take the product of Bass–Serre trees dual to a finite collection of filling curves), or products thereof. In [DJ99] Dranishnikov and Januszkiewicz show that any Coxeter group admits such an action on a finite product of trees. In particular, the same is true for any undistorted finitely generated subgroup, and also for any commensurable group (see below), and in particular it holds for right angled Artin groups. It then also follows for the Haglund–Wise virtually special groups, since these are commensurable to finitely generated undistorted subgroups of RAAGs.

The goal of this paper is to prove the following two theorems as an application of the projection complex techniques developed in [BBF15]; see also [BBFS19].

THEOREM 1.1. — *Let G be a residually finite hyperbolic group. Then G has (QT).*

THEOREM 1.2. — *Mapping class groups have (QT).*

Hamenstädt announced Theorem 1.2 in the Fall 2016, but our proof is different. Earlier, Hume [Hum17] constructed a (nonequivariant) quasi-isometric embedding of mapping class groups in a finite products of trees. In the Spring 2018 Hamenstädt also announced that Artin groups of finite type have (QT).

Cocompact lattices in $Sp(n,1), n > 1$ satisfy the assumptions of Theorem 1.1 and they have Kazhdan's property (T). In particular they do not have the Haagerup property, namely, they do not act properly by isometries on the Hilbert space. Recall also that if a group with property (T) acts on a tree, then it must have a fixed point (by Serre and Watatani, cf. [BdlHV08, Section 2.3]). On the other hand if a finitely generated group acts properly on a finite dimensional CAT(0) cube complex, e.g., a finite product of simplicial trees, then it has the Haagerup property, [NR97].

In view of the lattice example in $Sp(n,1)$, by Theorem 1.1, having a proper action on a finite product of quasi-trees that gives a quasi-isometric embedding of a group is not enough to expect a proper isometric group action on the Hilbert space. It is unknown if mapping class groups have either property (T) or the Haagerup property.

Property (QT) is a strong form of finiteness of asymptotic dimension. It was proved by Gromov [Gro93] that hyperbolic groups have finite asymptotic dimension, and by the authors in [BBF15] that mapping class groups do as well. See also [BHS17] for a quadratic bound. Also, it was known that a hyperbolic group admits a quasi-isometric embedding into the product of $n+1$ binary trees, where n is the topological dimension of the boundary at infinity of the group, [BDS07].

Finally, we remark that higher rank lattices do not have (QT) even though they have finite asymptotic dimension: by [FL08] an isometric action on a finite product of quasi-trees preserves the de Rham decomposition, and by Haettel [Hae20] higher rank lattices do not have non-elementary actions on quasi-trees.

We would like to thank the referee for comments, which improved the presentation of the paper.

2. Background with complements

2.1. Separability

We thank Chris Leininger and Ben McReynolds for pointing out the following fact. Recall that a subgroup $H < G$ is *separable* if it is the intersection of all finite index subgroups that contain it. Thus G is residually finite if and only if the trivial subgroup is separable.

LEMMA 2.1. — *Suppose G is residually finite. Then for every element $x \in G$ the centralizer*

$$C_G(x) = \{g \in G | gx = xg\}$$

is separable.

Proof. — Let $g \in G \setminus C_G(x)$, so $gxg^{-1}x^{-1} \neq 1$. We need to find a finite index subgroup $G' < G$ such that $g \notin G'$ but $G' \supset C_G(x)$. By residual finiteness there is a finite quotient $\overline{G}$ of G such that the above commutator maps nontrivially, i.e. the images $\overline{x}, \overline{g}$ of x, g do not commute. Then let G' be the preimage of $C_{\overline{G}}(\overline{x})$. □

We also note the following: *if $H < G$ is separable and the double coset HgH is distinct from H (i.e. $g \notin H$) then there is a finite index subgroup $G' < G$ disjoint from HgH.* Indeed, take G' so that $g \notin G' \supset H$.

2.2. Induction

We observe:
If $H < G$ has finite index, and H satisfies (QT), then so does G.
More generally, if H acts by isometries on a metric space X with orbit maps $H \to X$ QI embeddings, then G isometrically acts on $X^{[G:H]}$ with QI orbit maps. This is seen by the standard induction construction. Define

$$Y = Map_H(G, X)$$

as the set of H-equivariant functions $G \to X$ where H acts on G by left multiplication (and on X on the left). Then Y is a G-set via

$$(g \circ f)(\gamma) = f\left(\gamma g^{-1}\right).$$

Finally, as a metric space, Y is isometric to $X^{[G:H]}$: choose coset representatives g_i so that $G = \coprod Hg_i$ and define an isometry $Y \to X^{[G:H]}$ via

$$f \mapsto \left(f(g_1), f(g_2), \cdots, f(g_{[G:H]})\right).$$

2.3. Projection complexes

In this section we review the construction of projection complexes in [BBF15] with improvements from [BBFS19].

The input is a collection $\mathbf{Y}$ of geodesic metric spaces and for $X, Z \in \mathbf{Y}$ with $X \neq Z$ there is a projection $\pi_Z(X) \subset Z$. We also define $d_Y(X, Z) = \operatorname{diam} \pi_Y(X) \cup \pi_Y(Z)$ for $X, Y, Z \in \mathbf{Y}$. The pair $(\mathbf{Y}, \{\pi_Y\})$ satisfies the *projection axioms* for a *projection constant* $\xi \geqslant 0$ if

(P0) $\operatorname{diam} \pi_Y(X) \leqslant \xi$ when $X \neq Y$, (Bounded projection)
(P1) if X, Y, Z are distinct and $d_Y(X, Z) > \xi$ then $d_X(Y, Z) \leqslant \xi$, (Behrstock inequality)
(P2) for $X \neq Z$ the set
$$\{Y \in \mathbf{Y} | d_Y(X, Z) > \xi\}$$
is finite. (Finiteness)

If we replace (P1) with

(P1)′ if X, Y, Z are distinct and $d'_Y(X, Z) > \xi$ then $\pi'_X(Y) = \pi'_X(Z)$

then the collection satisfies the *strong projection axioms*. While there are many natural situations where the projection axioms hold, the strong projection axioms are not as natural. However, we can modify the projections so that they do hold. The following is proved in [BBFS19, Theorem 4.1 and Lemma 4.13].

THEOREM 2.2. — *If the collection $(\mathbf{Y}, \{\pi_Y\})$ satisfies the projection axioms then there are projections $\{\pi'_Y\}$ such that $(\mathbf{Y}, \{\pi'_Y\})$ satisfy the strong projections axioms with projection constant ξ' where $\pi'_Y(X)$ and $\pi_Y(X)$ are a uniform Hausdorff distance apart and ξ' only depends on ξ.*

Let $\mathcal{C}_K(\mathbf{Y})$ denote the space obtained from the disjoint union

$$\coprod_{Y \in \mathbf{Y}} Y$$

by joining points in $\pi_X(Z)$ with points in $\pi_Z(X)$ by an edge of length one whenever $d_Y(X, Z) < K$ for all $Y \in \mathbf{Y} \setminus \{X, Z\}$. When the spaces are graphs and projections are subgraphs we can join just the vertices in these projections. If a group G acts on the disjoint union of $Y \in \mathbf{Y}$ by isometries and the π_Y are G-invariant, ie, $\pi_{gY}(gX) = g\pi_Y(X)$, then G acts isometrically on $\mathcal{C}_K(\mathbf{Y})$.

THEOREM 2.3 ([BBF15]). — *If $(\mathbf{Y}, \{\pi_Y\})$ satisfy the strong projection axioms with projection constant ξ then for all $K > 2\xi$*

- *$\mathcal{C}_K(\mathbf{Y})$ is hyperbolic if all $Y \in \mathbf{Y}$ are δ-hyperbolic;*
- *$\mathcal{C}_K(\mathbf{Y})$ is a quasi-tree if all $Y \in \mathbf{Y}$ are quasi-trees with uniform QI constants.*

There is a very useful distance formula in $\mathcal{C}_K(\mathbf{Y})$, see [BBFS19, Theorem 6.3]. Let $X, Z \in \mathbf{Y}$ and $x \in X$, $z \in Z$. We define $d_Y(x, z) = d_Y(X, Z)$ if $Y \neq X, Z$, $d_X(x, z) = \operatorname{diam}(\{x\} \cup \pi_X(Z))$ if $X \neq Z$, and $d_X(x, z)$ is the given distance in X if $X = Z$. We also define the *distance function with threshold K* by

$$d_Y(,)_K = \begin{cases} d_Y(,) & \text{if } d_Y(,) \geqslant K \\ 0 & \text{otherwise.} \end{cases}$$

PROPOSITION 2.4. — *Let* $(\mathbf{Y}, \{\pi_Y\})$ *satisfy the strong projection axioms with projection constant* ξ. *Let* $x \in X$ *and* $z \in Z$ *be two points of* $\mathcal{C}(\mathbf{Y})$ *with* $X, Z \in \mathbf{Y}$. *Then*

$$\frac{1}{4} \sum_{Y \in \mathbf{Y}} d_Y(x, z)_K \leqslant d_{\mathcal{C}_K(\mathbf{Y})}(x, z) \leqslant 2 \sum_{Y \in \mathbf{Y}} d_Y(x, z)_K + 3K$$

for all $K \geqslant 4\xi$.

3. Proof of Theorem 1.1

For simplicity all metric spaces will be graphs with each edge of length 1 (and subspaces will be subgraphs).

3.1. Projection axioms in δ-hyperbolic spaces

PROPOSITION 3.1. — *Let* $\mathbf{Y}$ *be a collection of quasi-convex subspaces (with uniform constants) in a* δ-*hyperbolic space* $\mathcal{Y}$. *For* $X, Y \in \mathbf{Y}$ *let* $\pi_Y(X)$ *be the nearest point projection and assume that* $\operatorname{diam} \pi_Y(X) \leqslant \theta$ *(for all* $X \neq Y \in \mathbf{Y}$*). Then* $(\mathbf{Y}, \{\pi_Y\})$ *satisfies the projection axioms with projection constant* ξ.

Proof. — Axiom (P0) holds with constant θ by assumption. Given X and Z in $\mathbf{Y}$ let γ be a shortest geodesic from X to Z. Then for any other $Y \in \mathbf{Y}$ the nearest point projection of Y to γ will have diameter uniformly close to $d_Y(X, Z)$. To see this let α and β be shortest paths between X and Y and between Y and Z, respectively. The right endpoint of α will lie in $\pi_Y(X)$ and the left endpoint of β will lie in $\pi_Y(Z)$. Let γ' be a shortest path connecting these endpoints. If γ' is sufficiently long, when we concatenate these three geodesics we get a quasi-geodesic with coarsely the same endpoints as γ so it will fellow travel γ. By construction the nearest point projection of Y to the quasi-geodesic will be coarsely γ' (whose diameter is roughly $d_Y(X, Z)$) so the nearest point projection of Y to γ will also roughly be $d_Y(X, Z)$.

This directly implies (P1) since the distance from Z to α will be coarsely bounded below by $d_Y(X, Z)$ and hence, if this term is large, the strongly contracting property of δ-hyperbolic space implies that the projection of Z to α has uniformly bounded diameter. By the above assertion this diameter is coarsely $d_Z(X, Y)$ so this quantity is also bounded.

For (P2) if $X, Z, Y_1, \ldots, Y_n$ are all in $\mathbf{Y}$ with $d_{Y_i}(X, Z)$ large then the sum of the projections is bounded by, say, twice the distance between X and Z. If not there would be a Y_i and Y_j whose projection to γ have large diameter overlap which would imply that $\operatorname{diam} \pi_{Y_i}(Y_j) > \theta$. □

Quasi-geodesics

A quasi-geodesic is a subspace of a metric space that is quasi-isometric to $\mathbb{Z}$. For our purposes it will be convenient to assume that quasi-geodesics are a collection of bi-infinite paths parameterized by arc length. In the proof of Theorem 1.1 our

quasi-geodesics will be a single bi-infinite path. However, when we discuss mapping class groups we will need quasi-geodesics that are finite union of paths.

We now prove a sequence of technical results that will be needed in what follows. We begin with a general setup:

- $\mathcal{Y}$ is a δ-hyperbolic geodesic metric space. For convenience will assume that $\mathcal{Y}$ is a metric graph with edges of length one.
- $\widetilde{\mathcal{A}}$ is a collection of quasi-geodesics in $\mathcal{Y}$ with uniform constants.
- $\mathcal{A} \subset \widetilde{\mathcal{A}}$ is a sub-collection.
- For each distinct $\alpha, \beta \in \mathcal{A}$ the projection $\pi_\alpha(\beta)$ is a subset of α that is uniformly close (in the Hausdorff metric) to the nearest point projection.

We will refer to δ, the quasi-geodesic constants and the Hausdorff bound on the distance of the projections from the nearest point projections as the *coarse constants*.

The following is a direct consequence of Proposition 3.1.

THEOREM 3.2. — *Fix θ. Then there exists ξ, depending only on the coarse constants and θ, such that if $\operatorname{diam} \pi_\alpha(\beta) \leqslant \theta$ for all distinct α and β in $\mathcal{A}$ then $(\mathcal{A}, \{\pi_\gamma\})$ satisfies the projection axioms with projection constant ξ.*

The following proposition is the main estimate we need to approximate lengths in $\mathcal{Y}$ using our quasi-trees.

PROPOSITION 3.3. — *Fix constants $R, K > 0$. Then there exists an $L > 0$, depending only on the coarse constants and K, such that the following holds. Assume that*

- *any path of length L in some $\alpha \in \widetilde{\mathcal{A}}$ is contained in some $\gamma \in \mathcal{A}$;*
- *$\mathcal{A}$ is partitioned into $\mathcal{A}_1 \sqcup \ldots \sqcup \mathcal{A}_n$;*
- *for all $x, y \in \mathcal{Y}$ there is an $\alpha \in \widetilde{\mathcal{A}}$ that intersects the R-neighborhood of both x and y;*
- *$\widehat{x} = \{x_1, \ldots, x_n\}$ and $\widehat{y} = \{y_1, \ldots, y_n\}$ are n-tuples of vertices in $\mathcal{Y}$ that are contained in the R-neighborhoods of x and y, respectively.*

Then

$$d_{\mathcal{Y}}(x, y) \leqslant 2 \sum_i \sum_{\gamma \in \mathcal{A}_i} d_\gamma(x_i, y_i)_K + L + 2R.$$

Proof. — We can assume that $d_{\mathcal{Y}}(x, y) \geqslant 2R$. Choose an $\alpha \in \widetilde{\mathcal{A}}$ that intersects the R-neighborhood of both x and y. Let $\widetilde{\alpha}$ be the subpath of α between x and y that is disjoint from the R-neighborhoods of x and y but whose endpoints are exactly R from x and y. As geodesics (and hence quasi-geodesics) are strongly contracting in a δ-hyperbolic space, the projection of the R-neighborhood of x to any subpath of $\widetilde{\alpha}$ will be contained in a uniformly bounded neighborhood of the endpoint of the path closest to x (with the equivalent statement holding for the R-neighborhood of y). Therefore, for all $\beta \in \mathcal{A}$ that intersect $\widetilde{\alpha}$ and all $x', y' \in \mathcal{Y}$ with $d_{\mathcal{Y}}(x, x'), d_{\mathcal{Y}}(y, y') \leqslant R$, we have that

$$\operatorname{diam}(\widetilde{\alpha} \cap \beta) - d_\beta(x', y')$$

will be bounded above by a constant that only depends only on the coarse constants (and not on R). Using that quasi-geodesics in $\widetilde{\mathcal{A}}$ have uniform constants, if $\widetilde{\alpha} \cap \beta$

contains a subpath of sufficient path length then $\mathrm{diam}(\widetilde{\alpha} \cap \beta)$ will be large. When the diameter is large we can absorb the above additive error into a multiplicative one. Therefore there exists an $L > 0$ such that if $\widetilde{\alpha} \cap \beta$ contains a path of length L then

- $\mathrm{diam}(\widetilde{\alpha} \cap \beta) \leqslant 2d_\beta(x', y')$ and
- $\mathrm{diam}(\widetilde{\alpha} \cap \beta) \geqslant 2K$.

Combining these estimates we have

$$\mathrm{diam}(\widetilde{\alpha} \cap \beta) \leqslant 2d_\beta(x, y)_K$$

if $\widetilde{\alpha} \cap \beta$ contains a path of length L.

If $d_{\mathcal{Y}}(x, y) \geqslant L + 2R$ then $\widetilde{\alpha}$ will be a path of length at least L and by the choice of $\mathcal{A}$ we can find distinct axis $\gamma_1, \ldots, \gamma_m$ in $\mathcal{A}$ such that each $\gamma_i \cap \widetilde{\alpha}$ contains a segment of length L and the union of the intersections is all of $\widetilde{\alpha}$. We let $\mathcal{A}_{j_i}$ be the subcollection in the partition of $\mathcal{A}$ that contains γ_i. Using the above estimate we then have

$$\begin{aligned} d_{\mathcal{Y}}(x, y) &\leqslant \sum_i \mathrm{diam}(\widetilde{\alpha} \cap \gamma_i) + 2R \\ &\leqslant 2\sum_i d_{\gamma_i}(x_{j_i}, y_{j_i})_K + 2R \\ &\leqslant 2\sum_i \sum_{\gamma \in \mathcal{A}_i} d_\gamma(x_i, y_i)_K + 2R. \end{aligned}$$

If $d_{\mathcal{Y}}(x, y) < L + 2R$ then the sum in the inequality may be zero. However, if we add L to the right then the inequality will still hold in this case completing the estimate. $\qquad\square$

We now assume that G acts isometrically on $\mathcal{Y}$ and that $\widetilde{\mathcal{A}}$ is G-invariant. The action is *acylindrical* if for any $\epsilon > 0$ there exists $D, B > 0$ such that if $x, y, x', y' \in \mathcal{Y}$ with $d_{\mathcal{Y}}(x, y) > D$ then the set

$$\{g \in G | d_{\mathcal{Y}}(x', gx), d_{\mathcal{Y}}(y', gy) \leqslant \epsilon\}$$

has at most B elements. This is slightly different than the usual definition where one assumes that $x = x'$ and $y = y'$. It is not hard to check that the two definitions are equivalent.

In the proof of Theorem 1.2 we will consider the action of the mapping class group on the curve graphs of essential subsurfaces. For this reason we will need to consider actions where there is a large kernel. In particular assume that $\widetilde{G}$ acts on $\mathcal{Y}$ and G is the image of $\widetilde{G}$ in the isometry group of $\mathcal{Y}$. If the kernel of the quotient map $\widetilde{G} \to G$ is infinite then the action of $\widetilde{G}$ cannot be acylindrical. However, G may act acylindrically on $\mathcal{Y}$ in which case we say that the action of $\widetilde{G}$ has *acylindrical image*.

Let $\gamma \in \widetilde{\mathcal{A}}$ be the axis of an element g that acts hyperbolically on $\mathcal{Y}$.

We let $\widetilde{C}(\gamma)$ be the subgroup of $\widetilde{G}$ that fixes γ, up to bounded Hausdorff distance, and $\mathcal{A}_\gamma$ the equivalence classes, with respect to bounded Hausdorff distance, of the $\widetilde{G}$-translates of γ. There is a natural bijection between the set of left cosets of $\widetilde{C}(\gamma)$ and $\mathcal{A}_\gamma$. The group $\widetilde{G}$ acts on $\mathcal{A}_\gamma$. Accordingly, we need to define (the diameter of)

the projection between the equivalence classes:

$$\operatorname{diam} \pi_{[\gamma]}([\beta]) = \sup_{\gamma' \in [\gamma], \beta' \in [\beta]} \operatorname{diam} \pi_{\gamma'}(\beta').$$

Since the quasi-geodesic constants are uniformly bounded, the difference between $\operatorname{diam} \pi_{[\gamma]}([\beta])$ and $\operatorname{diam} \pi_{\gamma}(\beta)$ is uniformly bounded.

It will be convenient to assume that if $h \in \widetilde{G}$ and γ and $h(\gamma)$ are a bounded Hausdorff distance from each other then $h(\gamma) = \gamma$. So, we will work with this assumption, rather than work with the equivalence classes and the modified projection, which could be done with bounded modification of the constants in the argument.

If $g \in \widetilde{C}(\gamma)$ then, as g fixes γ (under our assumption), for any $\beta \in \mathcal{A}_\gamma$ we have that $\operatorname{diam} \pi_{\gamma}(\beta) = \operatorname{diam} \pi_{\gamma}(g(\beta))$. In particular, any two axes that are translates of γ by elements in the same double coset of $\widetilde{C}(\gamma)$ have projections to γ with the same diameter.

PROPOSITION 3.4. — *If $\widetilde{G}$ acts on $\mathcal{Y}$ with acylindrical image then there exists a $\theta > 0$, depending only on the coarse constants and the acylindrical constants, such that only finitely many double cosets of $\widetilde{C}(\gamma)$ have projection to γ of diameter $> \theta$.*

Proof. — First we can replace $\widetilde{G}$ with its image G in the isometry group of $\mathcal{Y}$. This is because the subgroup $C(\gamma)$ of G that fixes γ will be the image of $\widetilde{C}(\gamma)$ under the quotient map $\widetilde{G} \to G$ and the kernel of this quotient map will be also be the kernel of the quotient map $\widetilde{C}(\gamma) \to C(\gamma)$. Therefore the quotient map $\widetilde{G} \to G$ induces a bijection between the double cosets of $\widetilde{C}(\gamma)$ in $\widetilde{G}$ and of $C(\gamma)$ in G.

There is an $\epsilon > 0$, only depending on the coarse constants, such that if $\alpha, \beta \in \mathcal{A}$ then the difference between the diameter of $\pi_\alpha(\beta)$ and the diameter of the intersection of β with the ϵ-neighborhood of α is uniformly bounded. Let $D >> \epsilon$ be the acylindricity constant for ϵ. Let $\tilde{\gamma}$ be a finite subpath whose diameter is at least $4D$ and that contains at least two copies of a fundamental domain for the $C(\gamma)$ action on γ.

Assume that for $g \in C(\gamma)hC(\gamma)$ the translate $g(\gamma)$ has large projection to γ where "large" roughly means at least $2D$. Then we can assume that the coset representative h has been chosen such that there is subpath γ_h of $\tilde{\gamma}$ such that $h(\gamma_h)$ is contained in the ϵ-neighborhood of $\tilde{\gamma}$ and $\operatorname{diam} \gamma_h \geqslant 2D$. This implies that the endpoints of $h(\gamma_h)$ will be contained in the ϵ-neighborhood of two vertices x_h and y_h of $\tilde{\gamma}$ with $d_{\mathcal{Y}}(x_h, y_h) \geqslant D$. Note that there are finitely many triples $\{x, y, \alpha\}$ where $x, y \in \tilde{\gamma}$ with $d_{\mathcal{Y}}(x, y) \geqslant D$ and α is a subpath of $\tilde{\gamma}$ of diameter $\geqslant 2D$. By acylindricity for each triple $\{x, y, \alpha\}$ there are finitely many h such that $x_h = x$, $y_h = y$ and $\gamma_h = \alpha$. This implies that there are finitely many double cosets with projection roughly larger than $2D$. □

3.2. Axes

By the induction in Section 2.2 we may replace G by a finite index subgroup. Thus by residual finiteness we may assume G is torsion free (recall that hyperbolic groups contain finitely many conjugacy classes of torsion elements). In particular, if $\langle g \rangle$ is a

maximal cyclic subgroup, then the centralizer (and also normalizer) of g is $\langle g \rangle$ itself, which is therefore separable by Lemma 2.1.

The following is surely well known. We summarize the proof.

THEOREM 3.5. — *Let G be a torsion free δ-hyperbolic group and $\Gamma(G)$ a Cayley graph for some finite generating set. Then there exists a G-invariant collection $\widetilde{\mathcal{A}}$ of axes of maximal cyclic subgroups where the axes are uniform quasi-geodesics. Furthermore any x and y are within uniform distance R of an axis in $\widetilde{\mathcal{A}}$.*

Proof. — Let $|g|$ be the word norm with respect to the chosen generating set. In each conjugacy class of maximal cyclic subgroups choose a representative $\langle g \rangle$ with $|g|$ minimal possible. Define the axis γ_g as the union of g-translates of a geodesic segment from 1 to g and we assume that $\gamma_g = \gamma_{g^{-1}}$. We then extend the definition to the conjugates: $\gamma_{aga^{-1}} = a\gamma_g$ (this is well-defined by the remark about normalizers). These are axes of indivisible elements; each g acts by translation on its axis. Let $\widetilde{\mathcal{A}}$ be the collection of all such axes. Moreover, it is a well-known fact that each axis is a quasi-geodesic with uniform constants. Indeed, if $|g|$ is large compared to δ, say, $|g| > 1000\delta$ then there are uniform constants depending only on δ by [Gro87, 7.2C]. Now, there are only finitely many elements g with $|g| \leqslant 1000\delta$, so the claim follows. It is also standard that there exists a constant R such that for any two elements $x, y \in G$ there exists an axis γ_g that intersects the R-balls centered at x and y. This is a consequence of the fact [Gro87, 8.2G] that the set of pairs $(\gamma_g^\infty, \gamma_g^{-\infty}) \in \partial\Gamma(G) \times \partial\Gamma(G)$ for all g of infinite order is dense in $\partial\Gamma(G) \times \partial\Gamma(G)$. $\square$

3.3. Constants

We can now fix constants. The action of a group on its Cayley graph is proper and therefore acylindrical. By Proposition 3.4 there is a $\theta > 0$ such that for any axis $\gamma \in \widetilde{\mathcal{A}}$ there are only finitely many double cosets of $C(\gamma)$ that have projection to γ with diameter $> \theta$. By Theorem 3.2, there exist a $\xi' > 0$ such that for any subcollection of $\widetilde{\mathcal{A}}$ where the projections have diameter bounded by θ, the subcollection satisfies the projection axioms with projection constant ξ'. By Theorem 2.2 the projections can be modified to satisfy the strong projection axioms with projection constant ξ only depending ξ'. We then let $K = 4\xi$ so that the distance formula, Proposition 2.4, holds with threshold K. We then fix the segment constant L to satisfy Proposition 3.3 for K.

3.4. Preferred axes

We now choose the G-finite and G-invariant collection of preferred axes $\mathcal{A}$ (or equivalently, conjugacy classes of indivisible elements). We view the axes in $\widetilde{\mathcal{A}}$ as a collection of bi-infinite words in the generators and for every word x of length $\leqslant L$ choose, if possible, an element $\gamma_x \in \widetilde{\mathcal{A}}$ such that x is a subword of γ_x. Then let $\mathcal{A}$ be the collection of G-orbits of the selected axes. Note that every such x will not necessarily be a subword for an axis in $\mathcal{A}$ even if x is a geodesic but every subword x of length $\leqslant L$ in an axis in $\gamma \in \widetilde{\mathcal{A}}$ will be contained in an axis $\beta \in \mathcal{A}$ with $x \subset \gamma \cap \beta$.

3.5. Coloring $\mathcal{A}$

Let γ_1, ..., γ_n represent the distinct G-orbits of axes in $\mathcal{A}$. Then for each γ_i, $C(\gamma_i)$ is an infinite cyclic group and is its own centralizer. As G is residually finite, by Lemma 2.1 the subgroup $C(\gamma_i)$ is separable. Given $h \notin C(\gamma_i)$ there is a finite index subgroup of G that contains $C(\gamma_i)$ but not h and therefore doesn't contain the double coset $C(\gamma_i)hC(\gamma_i)$. Using Proposition 3.4 we can therefore find a finite index subgroup H_i such that the projection between any two axes in the H_i-orbit of γ_i (or the H_i-orbit of any axis in the G-orbit of γ_i) have diameter $\leqslant \theta$.

Let

$$H' = H_1 \cap \ldots \cap H_n$$

and let H be the intersection of the G-conjugates of H'. Now add axes γ_{n+1}, ..., γ_m so that we have one axis in $\mathcal{A}$ for each H-orbit. Let $\mathcal{A}_i$ be the H-orbit of γ_i. We then have:

COROLLARY 3.6. — *There is a finite index subgroup H of G and a partition $\mathcal{A}_1 \sqcup \ldots \sqcup \mathcal{A}_m$ of $\mathcal{A}$ such that each $\mathcal{A}_i$ is H-invariant and the projections between any two axes in a fixed $\mathcal{A}_i$ have diameter $\leqslant \theta$.*

3.6. Product of quasi-trees $\mathcal{X}$

By Theorem 2.3 and Proposition 2.4 for each $\mathcal{A}_i$ we have a quasi-tree $\mathcal{C}_K(\mathcal{A}_i)$ that has an isometric H-action and a lower bound on distance

$$\frac{1}{4} \sum_{\gamma \in \mathcal{A}_i} d_\gamma(x,y)_K \leqslant d_{\mathcal{C}_K(\mathcal{A}_i)}(x,y)$$

where x and y lie on axes in $\mathcal{A}_i$.

Let

$$\mathcal{X} = \prod_{i=1}^m \mathcal{C}_K(\mathcal{A}_i)$$

be the product of quasi-trees. We give $\mathcal{X}$ the ℓ^1-metric (which is quasi-isometric to the ℓ^2-metric). If $\widehat{x}$ and $\widehat{y}$ are m-tuples representing elements in $\mathcal{X}$ with the ith coordinate lying in axis in $\mathcal{A}_i$ then we sum the distance bound to get

$$\frac{1}{4} \sum_i \sum_{\gamma \in \mathcal{A}_i} d_\gamma(x_i,y_i)_K \leqslant d_{\mathcal{X}}(\widehat{x},\widehat{y})$$

Fix $\widehat{x}$ as a basepoint. We claim that the orbit map $H \to \mathcal{X}$ given by $h \mapsto h(\widehat{x})$ is a quasi-isometric embedding. As H is finite index in G it is quasi-isometrically embedded in $\Gamma(G)$ so we need to show that $d_{\mathcal{X}}(\widehat{x}, h(\widehat{x}))$ is bounded above and below by linear functions of the word length $|h|$. (We emphasize that the word length is for the generators of G we chose in Theorem 3.5.) The upper bound is clear since orbit maps are Lipschitz. The union $\widehat{x} \cup \{id\}$ is a finite set and therefore has diameter in $\Gamma(G)$ bounded by some $R > 0$. By Theorem 3.5, after possibly enlarging

R, we can also assume that for all $h \in H$ there is an axis in $\widetilde{A}$ that intersects the R-neighborhoods in $\Gamma(G)$ of both id and h. By Proposition 3.3 we have

$$|h| \leqslant 2 \sum_i \sum_{\gamma \in \mathcal{A}_i} d_\gamma(x_i, h(x_i))_K + L + 2R$$

and therefore

$$\frac{1}{8}(|h| - L - 2R) \leqslant d_{\mathcal{X}}(\widehat{x}, h(\widehat{x})).$$

This completes the proof of Theorem 1.1.

4. Proof of Theorem 1.2

Let Σ be a closed surface with finitely many marked points and let $MCG(\Sigma)$ be the mapping class group of Σ. The rest of the paper is devoted to the proof of Theorem 1.2, that $MCG(\Sigma)$ embeds in a product of quasi-trees. The general outline closely follows our proof of Theorem 1.1, but there are several complications that arise. The central one is that $MCG(\Sigma)$ is not a hyperbolic group. However, by the Masur–Minsky distance formula it does embed in an infinite product of hyperbolic spaces, the curve graphs for subsurfaces of Σ. In [BBF15], we used projection complexes to embed $MCG(\Sigma)$ in a finite product of hyperbolic spaces where the Masur–Minsky distance formula was a key ingredient. We would like to follow the strategy of the proof of Theorem 1.1 to embed each curve graph in finite product of quasi-trees. However, curve graphs are locally infinite so this adds a new difficulty.

To see this difficulty let us focus on the main factor, the curve graph $\mathcal{C}(\Sigma)$ of the surface Σ. If we mimic the construction in Section 3.4, we would take axes of all pseudo-Anosov elements of some bounded translation length. However, this would give us infinitely many conjugacy classes and the coloring construction in Section 3.5 will break down. To fix this problem we restrict to a finite collection of conjugacy classes that contain every *thick* segment of bounded length. This amounts to requiring the axes in Teichmüller space fellow travel every geodesic segment in a fixed thick part, but we will develop this notion combinatorially, in terms of Masur–Minsky subsurface projections. This will give an embedding of the thick part of the curve graph in a finite product of quasi-trees but not a quasi-isometric embedding of the entire curve graph. The distance lost will be picked up in curve graphs of proper subsurfaces. This is captured more formally in our *thick distance formula*, a version of the Masur–Minsky distance formula that counts only long segments that are thick in some subsurface. With these modifications, the proof of Theorem 1.1 will generalize to mapping class groups.

4.1. Curve graphs and subsurface projections

We set some notation. The *curve graph* of Σ is denoted $\mathcal{C}(\Sigma)$. If Y is an essential (connected, compact and possibly punctured) subsurface that is not a triply punctured sphere then its curve graph is also denoted $\mathcal{C}(Y)$. If x is a curve in $\mathcal{C}(\Sigma)$ then x is disjoint from Y if it can be homotoped in Σ to be disjoint from Y. Otherwise

x *cuts* Y. If x cuts Y we let $\pi_Y(x)$ be the projection of x to Y. If $\widetilde{x}$ is a collection of curves then $\pi_Y(\widetilde{x})$ is the union of $\pi_Y(x)$ for those $x \in \widetilde{x}$ that cut Y. If some component of ∂X cuts Y then $\pi_Y(X) = \pi_Y(\partial X)$. Two subsurfaces X and Y are *transverse* if a component of ∂X cuts Y and a component of ∂Y cuts X. We refer to [MM99], [MM00] for precise definitions.

The next result plays a central role in the paper.

THEOREM 4.1. — *There exists a universal constant $\xi > 0$ such that the following holds.*

- *If $\mathbf{Y}$ is a collection of pairwise transverse subsurfaces then $(\mathbf{Y}, \{\pi_Y\})$ satisfy the projection axioms with projection constant ξ (see [BBF15, Section 5]).*
- *If $x, z \in \mathcal{C}(\Sigma)$ and $d_Y(x,z) > \xi$ then every geodesic in $\mathcal{C}(\Sigma)$ from x to z contains a curve disjoint from Y ([MM00, Theorem 3.1]).*

These two results are usually stated separately but it will be convenient for us to have the same constant for both. The second bullet is the *Bounded Geodesic Image Theorem* and we will reference it below as BGIT. Sometimes the contrapositive will also be useful: If every curve in the geodesic cuts Y then the projection of the geodesic to Y has diameter $< \xi$ in $\mathcal{C}(Y)$.

4.2. The Masur-Minsky distance formula

Recall the Masur–Minsky distance formula for word length in the mapping class group ([MM00, Theorem 6.12], cf. [BBF15, Section 2] for this form).

THEOREM 4.2 (Masur–Minsky distance formula). — *Let $\widetilde{x}$ be a collection of filling curves on Σ. Then for R sufficiently large the word length $|g|$ (with respect to some fixed generating set) is bounded above and below by linear functions of*

$$\sum_{Y \subseteq \Sigma} d_Y(\widetilde{x}, g(\widetilde{x}))_R.$$

A collection of curves $\widetilde{x}$ is *filling* if every curve in $\mathcal{C}(\Sigma)$ intersects some curve in $\widetilde{x}$.

We will need a new version of this distance formula where length is only measured in the *thick* part of the curve graph. We need some more setup before we state the formula.

4.2.1. Bounded pairs and finiteness

The curve graph is not locally finite. The following concept is the replacement for this lack of local finiteness. See also [RS11].

DEFINITION 4.3 (*T*-thick). — *A collection of curves $\widetilde{x}$ is T-thick if for all $x, z \in \widetilde{x}$ and all proper subsurface Y we have $d_Y(x,z) \leqslant T$.*

THEOREM 4.4 ([CR07, Wat16]). — *Given any $C > 0$ there exists a $D > 0$ such that if x and y are in $\mathcal{C}(\Sigma)$ and $i(x,y) \geqslant D$ where $i(x,y)$ is the geometric intersection number, then $d_{\mathcal{C}(Y)}(x,y) \geqslant C$ for some subsurface $Y \subseteq \Sigma$.*

Up to the action of the mapping class group there are only finitely many curves of bounded intersection. This gives the following corollary.

COROLLARY 4.5. — *Up to the action of the mapping class group there are finitely many collections of T-thick curves in $\mathcal{C}(\Sigma)$ that have diameter in $\mathcal{C}(\Sigma)$ bounded by T.*

4.2.2. Tight geodesics

If both $x, z \in \mathcal{C}(\Sigma)$ cut Y, then we define
$$d_Y(x, z) = \operatorname{diam}_{\mathcal{C}(Y)}(\pi_Y(x) \cup \pi_Y(z)).$$
If g is a geodesic in $\mathcal{C}(\Sigma)$ connecting x to z and x' and z' are endpoints of a subsegment then there is no general relationship between $d_Y(x, z)$ and $d_Y(x', z')$. However, if we restrict to the special class of tight geodesics then we will get bounds. We say that a geodesic $g = \{x_0, x_1, \ldots, x_n\}$ is *tight* if x_i is a component of the boundary of the surface filled by x_{i-1} and x_{i+1} for $0 < i < n$. By [MM00, Lemma 4.5] there is a tight geodesic connecting any two curves in $\mathcal{C}(\Sigma)$.

LEMMA 4.6. — *Assume that x' and z' lie on a tight geodesic between x and z and that x' and z' cut a subsurface Y. If $d_Y(x', z') \geqslant \xi$ then x and z cut Y and*
$$d_Y(x', z') < d_Y(x, z) + 2\xi.$$
In particular if x and z are T-thick then the collection of curves in a tight geodesic from x to z is $(T + 2\xi)$-thick.

Proof. — Since $d_Y(x', z') \geqslant \xi$ by the BGIT there is a y' in between x' and z' that is disjoint from Y. If there is another $y \in g$ that is disjoint from Y and is in the complement of the segment between x' and y' then $d_{\mathcal{C}(\Sigma)}(y, y') \leqslant 2$ and therefore we must have that y and y' are both adjacent to either x' or y'. By tightness a curve that is adjacent to two curves that are disjoint from Y will also be disjoint from Y. This is a contradiction, so everything in the complement of the segment from x' to z' cuts Y. In particular x and z cut Y.

If $d_Y(x, x') \geqslant \xi$ there is a y between x and x' that is disjoint from Y, contradicting what we have just shown. Therefore $d_Y(x, x') < \xi$ and by the same argument $d_Y(z', z) < \xi$. By the triangle inequality
$$d_Y(x', z') \leqslant d_Y(x, z) + d_Y(x, x') + d_Y(z, z') < d_Y(x, z) + 2\xi. \qquad \square$$

Convention

Given a constant T let $\widehat{T} = T + 2\xi$ and $\widecheck{T} = T - 2\xi$.

4.2.3. Thick distance

Given filling collections $\widetilde{x}, \widetilde{z}$ on Σ and a subsurface $Y \subset \Sigma$, we define
$$\Omega_T(Y; \widetilde{x}, \widetilde{z}) = \{Z \subseteq Y \mid Z \neq Y, d_Z(\widetilde{x}, \widetilde{z}) > T\}.$$
This set has an order coming from inclusion. Let $\Omega_T^m(Y; \widetilde{x}, \widetilde{z})$ be the subset of maximal elements.

LEMMA 4.7. — *Given filling collections $\widetilde{x}, \widetilde{z}$ on Σ, and a subsurface $Z \subset \Sigma$ there are at most two subsurfaces Y with $d_Y(\widetilde{x}, \widetilde{z}) > T$ such that $Z \in \Omega_T^m(Y; \widetilde{x}, \widetilde{z})$.*

Proof. — Let Y_0, Y_1, Y_2 be subsurfaces such that $d_{Y_i}(\widetilde{x}, \widetilde{z}) > T$ and $Z \in \Omega_T^m(Y_i; \widetilde{x}, \widetilde{z})$. If $Y_i \subset Y_j$ for $i \neq j$ then $Y_i \in \Omega_T(Y_j; \widetilde{x}, \widetilde{z})$ so $Z \notin \Omega_T^m(Y_j; \widetilde{x}, \widetilde{z})$, a contradiction. Therefore the Y_i are mutually transverse.

Now choose an $x \in \pi_Y(\widetilde{x})$ and $z \in \pi_Y(\widetilde{z})$ that both cut Z (and hence the Y_i). By the ordering (see e.g. [BBF15, Theorem 3.3 (G)]) we have that two of the subsurfaces have a large projection to the third. We can assume that is Y_1 and $|d_{Y_1}(x,z) - d_{Y_1}(Y_0, Y_2)| \leqslant \xi$. In particular $d_{Y_1}(Y_0, Y_2)$ is large, so that ∂Y_0 and ∂Y_2 fill Y_1 and that if $Z \subset Y_1$ then it must intersect either ∂Y_0 or ∂Y_2, a contradiction. $\square$

We give a key definition.

DEFINITION 4.8 (T, R-thick distance). — *Fix sufficiently large constants T, R. Let $x_1, \ldots, x_n \in \mathcal{C}(Y)$ be curves occurring in this order on a tight geodesic in $\mathcal{C}(Y)$ from x to z such that $d_{\mathcal{C}(Y)}(x_{2i-1}, x_{2i}) \geqslant R$ and $d_Z(x_{2i-1}, x_{2i}) \leqslant T$ for all $Z \subsetneq Y$. Then the T, R-thick distance $d_Y^{T,R}(x,z)$ is the maximum of $\sum d_{\mathcal{C}(Y)}(x_{2i-1}, x_{2i})$ over all such choices for the x_i, and for the tight geodesics from x to z. For collections of curves $\widetilde{x}$ and $\widetilde{z}$ in $\mathcal{C}(Y)$ we set*

$$d_Y^{T,R}(\widetilde{x}, \widetilde{z}) = \max_{x \in \widetilde{x}, z \in \widetilde{z}} d_Y^{T,R}(x, z).$$

If $\widetilde{x}$ and $\widetilde{y}$ are collections in $\mathcal{C}(\Sigma)$ we define

$$d_Y^{T,R}(\widetilde{x}, \widetilde{z}) = d_Y^{T,R}\left(\pi_Y(\widetilde{x}), \pi_Y(\widetilde{z})\right).$$

DEFINITION 4.9 (Footprint). — *If g is a geodesic in $\mathcal{C}(Y)$ and $Z \subset Y$ is a proper subsurface then the footprint $F_Z(g)$ of Z is the set of vertices of g that are disjoint from Z. Since any vertices of $\mathcal{C}(Y)$ that are distance three or more apart will fill Y the diameter of $F_Z(g)$ is at most two. The footprint is connected for all $Z \subset \Sigma$ if and only if g is a tight geodesic.*

If $Z' \subset Z$ then $F_{Z'}(g) \supseteq F_Z(g)$ but it may be that strict inclusion holds. Whenever $F_Z(g)$ is nonempty we set $F_Z^{\subsetneq}(g)$ to be the union of $F_{Z'}(g)$ over all $Z' \subset Z$. Note that if y is in $F_Z^{\subsetneq}(g)$ then $d_{\mathcal{C}(Y)}(y, z) \leqslant 2$ for any boundary component z of Z. Thus the diameter of $F_Z^{\subsetneq}(g)$ will be at most four.

We state a key lemma.

LEMMA 4.10. — *For $T, R > 0$ sufficiently large the following holds. Let $\widetilde{x}$ and $\widetilde{z}$ be filling collections on Σ. Let Y be a subsurface in Σ. Then,*

$$d_Y^{T,R}(\widetilde{x}, \widetilde{z}) + (4 + 2R)\left|\Omega_T^m(Y; \widetilde{x}, \widetilde{z})\right| \geqslant d_Y(\widetilde{x}, \widetilde{z})_R$$

Proof. — First if $\Omega_T^m(Y; \widetilde{x}, \widetilde{z})$ is empty, then $\Omega_T(Y; \widetilde{x}, \widetilde{z})$ is empty and $d_Y^{T,R}(\widetilde{x}, \widetilde{y}) = d_Y(\widetilde{x}, \widetilde{y})_R$. So assume $\Omega_T^m(Y; x, z)$ is not empty.

Choose $x \in \pi_Y(\widetilde{x}), z \in \pi_Y(\widetilde{z})$ and g a tight geodesic between them such that g realizes the thick distance $d_Y^{T,R}(\widetilde{x}, \widetilde{z})$. Let J be a subsegment of g with endpoints x' and z'. By Lemma 4.6

$$d_Z(x', z') < d_Z(x, z) + 2\xi$$

if x' and y' cut Z. In particular if $d_Z(x', z') \geqslant T$ then

$$Z \in \Omega_{\check{T}}(Y; x, z) \subset \Omega_{\check{T}}(Y; \widetilde{x}, \widetilde{z}).$$

As every $Z \in \Omega_{\check{T}}(Y; \widetilde{x}, \widetilde{z})$ is contained in some $Z' \in \Omega_{\check{T}}^m(Y; \widetilde{x}, \widetilde{z})$ we have that if the interior of J is disjoint from every $Z' \in \Omega_{\check{T}}^m(Y; \widetilde{x}, \widetilde{z})$ then $d_Z(x', z') < T$ for all $Z \subset Y$.

Let $J_0, \ldots, J_n$ be a maximal collection of disjoint subsegments of g such that the interiors of the J_i do not contain any elements of $F_Z^\subseteq(g)$ for $Z \in \Omega_{\check{T}}^m(Y; \widetilde{x}, \widetilde{z})$. Then the endpoints of any J_i are T-thick. As each $F_Z^\subseteq(g)$ is connected $n \leqslant |\Omega_{\check{T}}^m(Y; x, z)|$.

Let

$$\mathcal{I} = \{i | 0 \leqslant i \leqslant n \text{ and } |J_i| \geqslant R\}$$

and $\mathcal{I}'$ the complement. Then

$$\sum_{0 \leqslant i \leqslant n} |J_i| = \sum_{i \in \mathcal{I}} |J_i| + \sum_{i \in \mathcal{I}'} |J_i| \leqslant d_Y^{T, R}(x, z) + R(|\Omega_{\check{T}}^m(Y; x, z)| + 1)$$

$$\leqslant d_Y^{T, R}(x, z) + 2R|\Omega_{\check{T}}^m(Y; x, z)|.$$

As the diameter of each $F_Z^m(g)$ is bounded above by four, the length of the complement of the J_i is bounded by $4|\Omega_{\check{T}}^m(Y; x, z)|$ so

$$d_Y(\widetilde{x}, \widetilde{z}) = d_{\mathcal{C}(Y)}(x, z)$$

$$\leqslant \sum |J_i| + 4 \left|\Omega_{\check{T}}^m(Y; x, z)\right|$$

$$\leqslant d_Y^{T, R}(x, z) + (4 + 2R) \left|\Omega_{\check{T}}^m(Y; x, z)\right|$$

and the Lemma 4.10 follows. $\qquad\square$

By $cx(Y)$ denote the complexity of Y, i.e. the length of the longest chain $Y = Y_0 \supset Y_1 \supset \cdots \supset Y_k$ of distinct subsurfaces.

THEOREM 4.11. — *Fix T, R sufficiently large with $R \leqslant \check{T}$. Let $\widetilde{x}, \widetilde{z}$ be filling collections in $\mathcal{C}(\Sigma)$. Then, for each n*

$$\sum_{cx(Y) \leqslant n} d_Y(\widetilde{x}, \widetilde{z})_{\check{T}} \leqslant \sum_{cx(Y) = n} d_Y^{T, R}(\widetilde{x}, \widetilde{z}) + (9 + 4R) \sum_{cx(Y) < n} d_Y(\widetilde{x}, \widetilde{z})_{\check{T}}.$$

We remark that each sum is over finitely many Y since it is for Y with $d_Y(\widetilde{x}, \widetilde{z}) \geqslant R$, and there are only finitely many such Y for given $\widetilde{x}, \widetilde{z}$.

Proof. — If $cx(Y) = n$ then by Lemma 4.10,

$$d_Y(\widetilde{x}, \widetilde{z})_{\check{T}} \leqslant d_Y(\widetilde{x}, \widetilde{z})_R$$

$$\leqslant d_Y^{T, R}(\widetilde{x}, \widetilde{z}) + (4 + 2R)|\Omega_{\check{T}}^m(Y; \widetilde{x}, \widetilde{z})|$$

$$\leqslant d_Y^{T, R}(\widetilde{x}, \widetilde{z}) + (4 + 2R) \sum_{Z \in \Omega_{\check{T}}^m(Y; \widetilde{x}, \widetilde{z})} d_Z(\widetilde{x}, \widetilde{z})_{\check{T}}.$$

By Lemma 4.7, any Z will appear in at most two $\Omega_{\check{T}}^m(Y; \widetilde{x}, \widetilde{z})$ and therefore if we sum the left hand side over all Y with $cx(Y) = n$ we have

$$\sum_{cx(Y) = n} d_Y(\widetilde{x}, \widetilde{z})_{\check{T}} \leqslant \sum_{cx(Y) = n} d_Y^{T, R}(\widetilde{x}, \widetilde{z}) + (8 + 4R) \sum_{cx(Y) < n} d_Y(\widetilde{x}, \widetilde{z})_{\check{T}}.$$

Adding $\sum_{cx(Y) < n} d_Y(\widetilde{x}, \widetilde{z})_{\check{T}}$ to both sides gives the inequality. $\qquad\square$

COROLLARY 4.12. — *Let $\widetilde{x}, \widetilde{z}$ be filling collections in $\mathcal{C}(\Sigma)$. Then for sufficiently large T, R with $R \leqslant \check{T}$,*

$$\sum_{Y \subset \Sigma} d_Y^{T,R}(\widetilde{x}, \widetilde{z}) \leqslant \sum_{Y \subset \Sigma} d_Y(\widetilde{x}, \widetilde{z})_R \leqslant (9 + 4R)^{cx(\Sigma)-1} \sum_{Y \subset \Sigma} d_Y^{T,R}(\widetilde{x}, \widetilde{z}).$$

Proof. — The first inequality is trivial since $d_Y^{T,R}(\widetilde{x}, \widetilde{z}) \leqslant d_Y(\widetilde{x}, \widetilde{z})_R$ for all Y. By inductively applying Theorem 4.11, with base case $n = cx(\Sigma)$, we have

$$\sum_{cx(Y) \leqslant cx(\Sigma)} d_Y(\widetilde{x}, \widetilde{z})_{\check{T}}$$

$$\leqslant (9 + 4R)^{cx(\Sigma)-n} \left(\sum_{n \leqslant cx(Y) \leqslant cx(\Sigma)} d_Y^{T,R}(\widetilde{x}, \widetilde{z}) + (9 + 4R) \sum_{cx(Y) < n} d_Y(\widetilde{x}, \widetilde{z})_{\check{T}} \right).$$

When $n = 1$ the last term on the right is zero. Since $\check{T} \leqslant R$ we have $d_R(\widetilde{x}, \widetilde{y}) \leqslant d_{\check{T}}(\widetilde{x}, \widetilde{y})$ and the result follows. □

4.2.4. Thick distance formula

Combining the Masur–Minsky distance formula (Theorem 4.2) with Corollary 4.12 we have our thick distance formula.

THEOREM 4.13 (Thick distance formula). — *Let $\widetilde{x}$ be a filling collection on Σ. Then for T, R sufficiently large with $R \leqslant \check{T}$, there exist C_0, C_1 such that for all $g \in MCG(\Sigma)$*

$$|g| \leqslant C_0 \sum_{Y \subseteq \Sigma} d_Y^{T,R}(\widetilde{x}, g(\widetilde{x})) + C_1$$

When we apply this result we will assume that $R = \check{T}$ and to simplify notation we set

$$d_Y^T(\widetilde{x}, \widetilde{y}) = d_Y^{T,\check{T}}(\widetilde{x}, \widetilde{y}).$$

4.3. Separability in the mapping class group

Let $\psi \in MCG(\Sigma)$ be a pseudo-Anosov. There are various equivalent characterizations, two of which are useful for us.

- ψ has positive stable translation length on $\mathcal{C}(\Sigma)$.
- ψ has positive translation length on the Teichmüller space $\mathcal{T}(\Sigma)$ with a unique invariant axis.

If ψ is a pseudo-Anosov then the orbit of any curve in $\mathcal{C}(\Sigma)$ will extend to a ψ-invariant quasi-geodesic γ and any two such invariant quasi-geodesics will be a bounded Hausdorff distance from each other. The *elementary closure*, $EC(\psi)$ is the subgroup of elements $\phi \in MCG(\Sigma)$ such that γ and $\phi(\gamma)$ are a bounded Hausdorff distance. Everything that commutes with ψ is contained in $EC(\psi)$ (including powers and roots) but there may be other elements.

The following is well known.

LEMMA 4.14. — *If $\psi \in MCG(\Sigma)$ is pseudo-Anosov then $EC(\psi)$ is virtually cyclic.*

Proof. — We use the second characterization of a pseudo-Anosov. Namely the action of ψ on the Teichmüller space $\mathcal{T}(\Sigma)$ has a unique axis and the subgroup $EC(\psi)$ will preserve the axis and fix its endpoints at infinity. Translation length along the axis will define a homomorphism to $\mathbb{R}$ with discrete image. The subgroup of $EC(\psi)$ of elements with translation length zero will fix the axis pointwise and will therefore be finite since the stabilizer of any element in $\mathcal{T}(\Sigma)$ will be finite. In particular there is a surjective map of $EC(\psi)$ to $\mathbb{Z}$ with finite kernel. The lemma follows. $\square$

We note that one can also prove this lemma by applying a general result ([DGO17, Lemma 6.5]) which says that if a group G acts on a hyperbolic space, then the elementary closure (with respect to this action) of an element $g \in G$ is virtually cyclic if g is "WPD" (weak proper discontinuous).

Let $Y \subset \Sigma$ be a proper subsurface and let $MCG(\Sigma; Y)$ be the subgroup of the mapping class group that preserves Y. If Y is non-annular let $\bar{Y}$ be the surface obtained by collapsing the components of ∂Y to marked points. There is a natural homomorphism

$$MCG(\Sigma; Y) \to MCG(\bar{Y}).$$

The kernel of this homomorphism are mapping classes that can be represented by homeomorphisms that are the identity on Y. Furthermore every mapping class in the image of the homomorphism is the image of a mapping class that is the identity on the complement of Y and these two types of mapping classes commute.

Given a $\psi \in MCG(\Sigma; Y)$ we say that ψ is *pseudo-Anosov on Y* if its image in $MCG(\bar{Y})$ is pseudo-Anosov. The *elementary closure with respect to Y*, $EC(\psi; Y)$, is the subgroup of elements $\phi \in MCG(\Sigma; Y)$ whose image in $MCG(\bar{Y})$ is contained in $EC(\bar{\psi})$. The image $\overline{EC}(\psi; Y)$ of $EC(\psi; Y)$ in $MCG(\bar{Y})$ is a subgroup of $EC(\bar{\psi})$. Note that image of $\overline{EC}(\psi; Y)$ will be infinite but it may be a proper subgroup of $EC(\bar{\psi})$. In particular, $\overline{EC}(\psi; Y)$ contains an infinite cyclic group.

LEMMA 4.15. — *Let Y be a non-annular subsurface and assume that $\psi \in MCG(\Sigma; Y)$ with image $\bar{\psi}$ in $MCG(\bar{\psi})$ pseudo-Anosov. Then $EC(\psi; Y)$ is a centralizer in $MCG(\Sigma)$.*

Proof. — Choose $\phi \in MCG(\Sigma; Y)$ such that ϕ is the identity on the complement of Y and its image $\bar{\phi}$ in $MCG(\bar{Y})$ is a primitive element of infinite order in $\overline{EC}(\psi; Y)$. By Lemma 4.14, $\overline{EC}(\psi; Y)$ is virtually cyclic so there is a short exact sequence

$$1 \to F \to \overline{EC}(\psi; Y) \to \mathbb{Z} \to 1$$

where F is finite and the subgroup $\langle \bar{\phi} \rangle \subset \overline{EC}(\psi; Y)$ surjects onto $\mathbb{Z}$. The subgroup $\langle \bar{\phi} \rangle$ acts on the finite group F by conjugation so there is a k such that conjugation by $\bar{\phi}^k$ is the identity on F. That is $\bar{\phi}^k$ commutes with every element of F. As every element of $\overline{EC}(\psi; Y)$ is a product of an element of F and a power of $\bar{\phi}$ this implies that $\bar{\phi}^k$ commutes with every element of $\overline{EC}(\psi; Y)$ and the centralizer of $\bar{\phi}^k$ in $\overline{EC}(\psi; Y)$ is the entire group.

We now claim that the centralizer of ϕ^k in $MCG(\Sigma)$ is $EC(\psi; Y)$. Any element that commutes with ϕ^k will be contained in $EC(\psi; Y)$ so we only need to show that every element of $EC(\psi; Y)$ commutes with ϕ^k. We can decompose every element of $EC(\psi; Y)$ as a composition of three elements:

- a mapping class ϕ_0 that is the identity on Y;
- a mapping class ϕ_1 that is the identity on the on the complement of Y and has finite image in $MCG(\bar{Y})$;
- a power of ϕ.

As ϕ^k will commute with both ϕ_0 and any power of ϕ we only need to show that ϕ^k commutes with ϕ_1. The image $\bar{\phi}_1$ of ϕ_1 in $MCG(\bar{Y})$ has finite order so there exists a ℓ such that $\bar{\phi}_1^\ell$ is the identity. Therefore ϕ_1^ℓ has a representative that is the identity on Y and therefore commutes with ϕ^k.

The image of the commutator $[\phi_1, \phi^k]$ in $MCG(\bar{Y})$ is $[\bar{\phi}_1, \bar{\phi}^k]$ and is trivial as these two elements commute. This implies that $[\phi_1, \phi^k]$ has a representative that is the identity on Y. But it also has a representative that is the identity on the complement of Y (as both ϕ_1 and ϕ^k do). This implies that $[\phi_1, \phi^k]$ is a composition of Dehn twists in ∂Y. In particular it is either trivial or of infinite order. As $[\phi_1, \phi^k]$ commutes with both ϕ_1 and ϕ^k a straightforward calculation shows that $[\phi_1^j, \phi^k] = [\phi_1, \phi^k]^j$ and therefore if $[\phi_1, \phi^k]$ is non-trivial then it is of infinite order. However, we observed above that ϕ_1^ℓ and ϕ^k commute and therefore $[\phi_1^\ell, \phi^k]$ is trivial. This implies that ϕ_1 and ϕ^k commute, completing the proof of Lemma 4.15. $\qquad\square$

We say that a quasi-geodesic $\gamma \subset \mathcal{C}(Y)$ is an *axis* if there is a $\psi \in MCG(\Sigma; Y)$ and γ is $EC(\psi; Y)$-invariant. Every ψ that is pseudo-Anosov on Y has an axis that can be obtained by taking the $\overline{EC}(\psi; Y)$ translates of a ψ-invariant bi-infinite path.

To match the notation from Theorem 1.1 we let $C(\gamma) \subset MCG(\Sigma)$ be the stabilizer of γ. When γ is an axis for ψ we have $C(\gamma) = EC(\psi; Y)$.

By [Gro75], the mapping class group is residually finite. Then Lemma 2.1 combined with Lemmas 4.15:

COROLLARY 4.16. — *If $\gamma \subset \mathcal{C}(Y)$ is an axis then $C(\gamma)$ is separable.*

Note that to this point we have not discussed the case when Y is an annulus. Here all of $\mathcal{C}(Y)$ is a quasi-geodesic and will play the role of an axis. While it is true that the stabilizer of $\mathcal{C}(Y)$ is separable ([LM07]) we will not use this.

4.4. Projection axioms for axes in curve graphs

We begin with our setup for mapping class groups:

- $\mathbf{Y}$ is a collection of transverse subsurfaces.
- $\widetilde{\mathcal{A}}_{\mathbf{Y}}$ is the collection of all quasi-geodesics in all the curve graphs $\mathcal{C}(Y), Y \in \mathbf{Y}$, with uniform constants. (We fix uniform quasi-geodesic constants to start with.) For each subsurface Y we let $\widetilde{\mathcal{A}}_Y$ be the subcollection contained in $\mathcal{C}(Y)$.
- $\mathcal{A}_{\mathbf{Y}} \subset \widetilde{\mathcal{A}}_{\mathbf{Y}}$ is a subcollection.

- For $\alpha \in \mathcal{A}_X$ and $\beta \in \mathcal{A}_X$ we define $\pi_\alpha(\beta)$ to be uniformly close (in the Hausdorff metric) to the nearest point projection of α to β.
- For $\alpha \in \mathcal{A}_X$ and $\beta \in \mathcal{A}_Z$ when $X \neq Z$ we define $\pi_\alpha(\beta)$ to be uniformly close to the nearest point projection of $\pi_X(Z)$ to α.

For mapping class groups the coarse constants are the quasi-geodesic constants above along with the projection constant and BGIT constant from Theorem 4.1 and the hyperbolicity constant for curve graphs.

Theorem 4.17. — *Fix $\theta > 0$. Then there exists $\chi > 0$, depending only on the coarse constants and θ, such that if $\operatorname{diam} \pi_\alpha(\beta) \leqslant \theta$ whenever α and β are distinct elements in the same $\mathcal{A}_Y$ then $(\mathcal{A}, \{\pi_\gamma\})$ satisfies the projection axioms with projection constant χ.*

Proof. — If $\gamma_0, \gamma_1 \in \mathcal{A}_Y$ then (P0) holds by assumption. If γ_0 and γ_1 are in distinct $\mathcal{A}_Y$ then (P0) holds by Theorem 4.1.

For the remaining two axioms we first observe:

($*$) If $\gamma_0, \gamma_1, \gamma_2$ are in $\mathcal{A}_{Y_0}, \mathcal{A}_{Y_1}, \mathcal{A}_{Y_2}$ and the Y_0 and Y_2 are distinct from Y_1 then

$$d_{\gamma_1}(\gamma_0, \gamma_2) \text{ is coarsely bounded above by } d_{Y_1}(Y_0, Y_2).$$

This follows from the fact that nearest point projections in δ-hyperbolic spaces are coarsely Lipschitz.

When all three Y_i are distinct then (P1) follows directly from ($*$) and Theorem 4.1. If the three Y_i are all equal then (P1) follows from Proposition 3.1. The last case is when $Y_0 = Y_1$ but they are distinct from Y_2. In this case $d_{\gamma_2}(\gamma_0, \gamma_1)$ will be uniformly bounded ($*$). Applying Proposition 3.1 to γ_0, γ_1 and $\pi_{Y_0 = Y_1}(Y_2)$ we see that at most one of $d_{\gamma_0}(\gamma_1, \gamma_2)$ and $d_{\gamma_1}(\gamma_0, \gamma_2)$ are large proving (P1) in this final case.

Now we prove (P2). Fix $\alpha \in \mathcal{A}_X$ and $\beta \in \mathcal{A}_Z$. If Y is distinct from X and Z and $\gamma \in \mathcal{A}_Y$ with $d_\gamma(\alpha, \beta)$ large then by ($*$) we have that $d_Y(X, Z)$ is large. Therefore Theorem 4.1 implies that there are finitely many Y such that $\mathcal{A}_Y$ contains a γ with $d_\gamma(\alpha, \beta)$ large. Applying Proposition 3.1 to the collection of quasi-convex sets $\mathcal{A}_Y \cup \{\pi_Y(X), \pi_Y(Z)\}$ we see that in each such Y there are finitely many $\gamma \in \mathcal{A}_Y$ with $d_\gamma(\alpha, \beta)$ large. Similarly we get finitely many $\gamma \in \mathcal{A}_X$ with $d_\gamma(\alpha, \beta)$ large by applying Proposition 3.1 to $\mathcal{A}_X \cup \{\pi_X(Z)\}$ if X and Z are distinct or simply to A_X if $X = Z$. This proves (P2). $\square$

Next we prove the version of Proposition 3.3 that we need for the mapping class group.

Proposition 4.18. — *Fix $K > 0$. Then there exists an $T > 0$ depending only on the coarse constants and K such that the following holds. Assume that*

- *$\tilde{x}, \tilde{y}$ are filling collections in $\mathcal{C}(\Sigma)$;*
- *$\hat{x} = (x_1, x_2, \cdots, x_n)$ and $\hat{y} = (y_1, y_2, \cdots, y_n)$ are n-tuples of curves with each x_i and y_i lying in curve graphs $\mathcal{C}(X_i)$ and $\mathcal{C}(Y_i)$ with X_i and Y_i in $\mathbf{Y}$;*
- *any tight geodesic segment in $\mathcal{C}(Y)$, with $Y \in \mathbf{Y}$, that is $\widehat{T}$-thick and of length $\check{T}$ is contained in some $\gamma \in \mathcal{A}_Y$;*
- *$\mathcal{A}_{\mathbf{Y}}$ is partitioned into $\mathcal{A}_1 \sqcup \cdots \sqcup \mathcal{A}_n$.*

Then

$$\sum_{Y \in \mathbf{Y}} d_Y^T(\widetilde{x}, \widetilde{y}) \leqslant 2 \sum_i \sum_{\gamma \in \mathcal{A}_i} d_\gamma(x_i, y_i)_K + \frac{2T}{\xi} \sum_{Y \in \mathbf{Y}} \left(d_Y(\widetilde{x}, \widehat{x})_\xi + d_Y(\widetilde{y}, \widehat{y})_\xi \right).$$

Proof. — We choose T in a way similar to the choice of L in the proof of Proposition 3.3: Fix $x, x', y, y' \in \mathcal{C}(Y)$ such that $d_{\mathcal{C}(Y)}(x, x') \leqslant d_Y(\widetilde{x}, \widehat{x})$, $d_{\mathcal{C}(Y)}(y, y') \leqslant d_Y(\widetilde{y}, \widehat{y})$ and

$$d_{\mathcal{C}(Y)}(x, y) > d_Y(\widetilde{x}, \widehat{x}) + d_Y(\widetilde{y}, \widehat{y}).$$

and let $\widetilde{\alpha}$ be the subsegment of a geodesic from x to y where the $\max\{d_Y(\widetilde{x}, \widehat{x}) - \xi, 0\}$ and $\max\{d_Y(\widetilde{y}, \widehat{y}) - \xi, 0\}$ neighborhoods of each endpoint have been removed. Note that $d_Y(\widetilde{x}, \widehat{x}) \leqslant \xi$ and $d_Y(\widetilde{y}, \widehat{y}) \leqslant \xi$ for all but finitely many Y. Then the nearest point projection of x and x' to any subsegment of $\widetilde{\alpha}$ will be in a uniformly bounded neighborhood of the endpoint closest to x with a similar statement for the projection of y and y'. As in Proposition 3.3, for any $\beta \in \mathcal{A}_\mathbf{Y}$ we have a uniform upper bound on

$$\operatorname{diam}(\widetilde{\alpha} \cap \beta) - d_\beta(x', y')$$

and we can choose T such that if $\widetilde{\alpha} \cap \beta$ contains a path of length $\check{T}$ then

- $\operatorname{diam}(\widetilde{\alpha} \cap \beta) \leqslant 2 d_\beta(x', y')$ and
- $\operatorname{diam}(\widetilde{\alpha} \cap \beta) \geqslant 2K$.

For each subsurface Y where $d_Y^T(\widetilde{x}, \widetilde{y}) > 0$ we let α_Y be a tight geodesic between $x \in \pi_Y(\widetilde{x})$ and $y \in \pi_Y(\widetilde{y})$ that realizes $d_Y^T(\widetilde{x}, \widetilde{y})$. In particular there are disjoint subsegments $\alpha_1^Y, \ldots, \alpha_{n_Y}^Y$ of α_Y such that endpoints of α_i^Y are T-thick and $d_Y^T(\widetilde{x}, \widetilde{y})$ is the sum of the lengths of the α_i^Y. By Lemma 4.6, each subsegment of length $\check{T}$ in each α_i^Y will have endpoints that are $\widehat{T}$-thick and therefore each such subsegment will be contained in $\gamma \in \mathcal{A}_\mathbf{Y}$. Let $\widetilde{\alpha}_Y$ be obtained by removing the $d_Y(\widetilde{x}, \widehat{x}) - \xi$ and $d_Y(\widetilde{y}, \widehat{y}) - \xi$ neighborhoods of each endpoint of α_Y. By the above estimate if $\gamma \in \mathcal{A}_i$ and $\widetilde{\alpha}_Y \cap \gamma$ contains a path of length $\check{T}$ then

$$K \leqslant \frac{1}{2} \operatorname{diam}(\widetilde{\alpha}_Y \cap \gamma) \leqslant d_\gamma(x_i, y_i).$$

Let $\mathcal{I}$ be the indices i such that $\operatorname{diam}(\widetilde{\alpha}_Y \cap \alpha_i^Y) \geqslant \check{T}$. Then

$$\sum_{i \in \mathcal{I}} \operatorname{diam}\left(\widetilde{\alpha}_Y \cap \alpha_i^Y\right) \leqslant 2 \sum_j \sum_{\gamma \in \mathcal{A}_j \cap \mathcal{A}_Y} d_\gamma(x_j, y_j)_K.$$

To complete the proof we will show that

$$d_Y^T(\widetilde{x}, \widetilde{y}) - \sum_{i \in \mathcal{I}} \operatorname{diam}\left(\widetilde{\alpha}_Y \cap \alpha_i^Y\right) \leqslant \frac{2T}{\xi} \left(d_Y(\widetilde{x}, \widehat{x})_\xi + d_Y(\widetilde{y}, \widehat{y})_\xi \right).$$

If $d_Y(\widetilde{x}, \widehat{x})_\xi = d_Y(\widetilde{y}, \widehat{y})_\xi = 0$ then $\alpha_Y = \widetilde{\alpha}_Y$ and the two terms are equal and the difference is zero proving the bound in this case. If both $d_Y(\widetilde{x}, \widehat{x}) \geqslant \xi$ and $d_Y(\widetilde{y}, \widehat{y}) \geqslant \xi$ then

$$d_Y^T(\widetilde{x}, \widetilde{y}) - \sum_{i \in \mathcal{I}} \operatorname{diam}\left(\widetilde{\alpha}_Y \cap \alpha_i^Y\right) \leqslant d_Y(\widetilde{x}, \widehat{x}) + d_Y(\widetilde{y}, \widehat{y}) - 2\xi + 2\check{T}.$$

Note that $2\check{T}$ on the right comes from the fact that there may be two (but no more) α_i^Y that intersect $\widetilde{\alpha}_Y$ in a segment of length $< \check{T}$. Similarly if only $d_Y(\widetilde{x}, \widehat{x}) \geqslant \xi$ then

$$d_Y^T(\widetilde{x}, \widetilde{y}) - \sum_{i \in \mathcal{I}} \operatorname{diam}\left(\widetilde{\alpha}_Y \cap \alpha_i^Y\right) \leqslant d_Y(\widetilde{x}, \widehat{x}) - \xi + \check{T}.$$

If only $d_Y(\widetilde{y}, \widehat{y}) \geqslant \xi$ the roles of x and y are swapped in the above inequality. If we combine these bounds with the fact that if $C \geqslant \xi$ then

$$C - \xi + \check{T} \leqslant \frac{2T}{\xi} C$$

we get the desired bound in all cases.

The proof is then completed by summing the inequality

$$d_Y^T(\widetilde{x}, \widetilde{y}) \leqslant 2 \sum_j \sum_{\gamma \in \mathcal{A}_j \cap \mathcal{A}_Y} d_\gamma(x_j, y_j)_K + \frac{2T}{\xi}\left(d_Y(\widetilde{x}, \widehat{x})_\xi + d_Y(\widetilde{y}, \widehat{y})_\xi\right)$$

over all $Y \in \mathbf{Y}$. $\qquad\square$

4.5. Axes

Here is our replacement for Theorem 3.5 in the setting of mapping class groups.

THEOREM 4.19. — *There exists a $MCG(\Sigma)$-invariant collection of uniform quasi-geodesics $\widetilde{\mathcal{A}}$ with $\widetilde{\mathcal{A}}_Y$ the subcollection of $\widetilde{\mathcal{A}}$ contained in $\mathcal{C}(Y)$ such that*
- *if Y is an annulus then $\widetilde{\mathcal{A}}_Y = \{\mathcal{C}(Y)\}$;*
- *if Y is non-annular then $\widetilde{\mathcal{A}}_Y$ is a collection of axes and every geodesic segment $\sigma \in \mathcal{C}(Y)$ of length $\geqslant 3$ is contained in some $\gamma \in \widetilde{\mathcal{A}}_Y$.*

We need a preliminary lemma.

LEMMA 4.20. — *Let $z_0, \ldots, z_n$ be a collection of curves in $\mathcal{C}(Y)$ such that*

$$d_{z_i}(z_{i-1}, z_{i+1}) \geqslant 3\xi$$

for $i = 1, \ldots, sn - 1$. Then

$$d_Y(z_0, z_n) \geqslant \sum_{i=1}^n d_{\mathcal{C}(\Sigma)}(z_{i-1}, z_i) + 2 - 2n.$$

Proof. — We first show that if $0 < i < n$ then

$$|d_{z_i}(z_0, z_n) - d_{z_i}(z_{i-1}, z_{i+1})| \leqslant 2\xi$$

and therefore

$$d_{z_i}(z_0, z_n) \geqslant d_{z_i}(z_{i-1}, z_{i+1}) - 2\xi \geqslant \xi.$$

We induct on n where the base case is when $n = 2$. Since $i < n$ we have

$$d_{z_{i-1}}(z_0, z_i) \geqslant \xi$$

so by (P1)

$$d_{z_i}(z_0, z_{i-1}) \leqslant \xi.$$

Similarly

$$d_{z_i}(z_{i+1}, z_n) \leqslant \xi$$

and the desired bounds follow from the triangle inequality. We note that this also proves that all of the z_i intersect.

We now prove the distance estimate via induction. The base case is when $n = 1$ and the inequality is an equality by observation. Now assume the estimate holds for k. By the BGIT any geodesic from z_0 to z_k must pass within one of z_k. Then by the triangle inequality

$$d_{\mathcal{C}(Y)}(z_0, z_{k+1}) \geqslant d_{\mathcal{C}(Y)}(z_0, z_k) + d_{\mathcal{C}(Y)}(z_k, z_{k+1}) - 2$$

and the bound follows. □

Proof of Theorem 4.19. — Let σ be a geodesic segment of length $\geqslant 3$ in some $\mathcal{C}(Y)$ with Y non-annular. Let x and y be the endpoints of σ. We show that for a sufficiently large positive integer n the composition of Dehn twists $\psi = D_y^n D_x^n$ is a pseudo-Anosov on Y with an axis (with uniform constants) that contains σ. Let $z_{2k} = \psi^k(x)$ and $z_{2k+1} = \psi^k(y)$. Let

$$\gamma = \cup_k \psi^k \left(\sigma \cup D_x^n(\sigma) \right).$$

This is a ψ-equivariant path where the z_k appear in the order given by their indices and the path is geodesic between each z_k and z_{k+1}. Note that

$$d_{z_i}(z_{i-1}, z_{i+1}) = d_{z_i} \left(z_{i-1}, D_{z_i}^n(z_{i-1}) \right) \sim n$$

so when n is sufficiently large the z_i satisfies the conditions of Lemma 4.20. The estimate there implies that γ is a quasi-geodesic with uniform constants. To get a quasi-geodesic that is $EC(\psi; Y)$-invariant we take γ_σ to be the $EC(\psi; Y)$ translates of γ. This is a collection (in fact a finite collection) of quasi-geodesics with uniform constants that are all in a bounded Hausdorff distance of each other. Therefore γ_σ is uniformly quasi-isometric to $\mathbb{Z}$.

The collection of geodesic segments in curve graphs $\mathcal{C}(Y)$ with Y non-annular is $MCG(\Sigma)$-invariant. We choose a representative in each $MCG(\Sigma)$-orbit and apply the above construction and then take the $MCG(\Sigma)$-orbit of this collection of axes. Finally we add all of the curve graphs $\mathcal{C}(Y)$ with Y annular to form $\widetilde{\mathcal{A}}$. □

4.6. Constants

As in the proof of Theorem 1.1 we are now ready to fix constants. Fix $\theta > 0$ as in Proposition 3.4 such that for any axis $\gamma \in \widetilde{\mathcal{A}}_Y$ there are only finitely many double cosets in of $C(\gamma)$ in $MCG(\Sigma; Y)$ whose projection to γ is $> \theta$. Choose $\chi' > 0$ to be the projection constant from Theorem 4.17 with diameter bound θ. As before we modify the projections so that the strong projection axioms hold with projection constant χ and let $K = 4\chi$ so that the distance formula holds with threshold K. When then choose T with respect to K as in Proposition 4.18.

4.7. Preferred axes

By Corollary 4.5 there are finitely many $\widehat{T}$-thick geodesic segments σ_1, ..., σ_n of length $\check{T}$ such that every other $\widehat{T}$-thick geodesic segment of length $\check{T}$ is contained in the $MCG(\Sigma)$-orbit of one of the σ_i. By Theorem 4.19, for each σ_i there exists a $\gamma_i \in \widetilde{\mathcal{A}}$ such that $\sigma_i \subset \gamma_i$. Let $\mathcal{A}$ be the $MCG(\Sigma)$-orbits of the γ_i.

4.8. Coloring $\mathcal{A}$

We can choose a subgroup $G < MCG(\Sigma)$ such that the G-orbit of a subsurface is a transverse collection. When Σ is a closed surface this is Theorem [BBF15, Lemma 5.7]. When Σ has punctures, the statement follows by blowing up the punctures to boundary components to obtain a surface $\widetilde{\Sigma}$ and doubling to obtain the surface $D\widetilde{\Sigma}$. Then $MCG(\widetilde{\Sigma})$ is a subgroup of $MCG(D\widetilde{\Sigma})$ and we have the desired finite index subgroup by taking the intersection with the one in $MCG(D\widetilde{\Sigma})$. Finally we project this subgroup to $MCG(\Sigma)$.

By Corollary 4.16 the stabilizer $C(\gamma_i)$ is separable in $MCG(\Sigma)$ and therefore in G. By our choice of θ from Proposition 3.4 we can use the separability $C(\gamma_i)$ to find a finite index subgroup H_i of G such that if $h \in H_i \cap MCG(\Sigma; Y)$ then $\operatorname{diam} \pi_{\gamma_i}(h(\gamma_i)) < \theta$. Let

$$H = H_1 \cap \ldots \cap H_n$$

and add axes γ_{n+1}, ..., γ_m so that we have one axis in $\mathcal{A}$ for each H-orbit. Let $\mathcal{A}_i$ be the H-orbit of γ_i. The $\mathcal{A}_i$ will partition $\mathcal{A}$ and each one will satisfy the conditions of Theorem 4.17.

4.9. Product of quasi-trees $\mathcal{X}$

Again we follow the proof of Theorem 1.1. For each $\mathcal{A}_i$ we have the quasi-tree $\mathcal{C}_K(\mathcal{A}_i)$. Choose a filling collection $\widetilde{x}$ and an m-tuple of curves $\widehat{x} = \{x_1, \ldots, x_m\}$ such that each x_i lies in some axis in $\mathcal{A}_i$. We let

$$\mathcal{X} = \prod_{i=1}^{m} \mathcal{C}_K(\mathcal{A}_i)$$

with the ℓ^1-metric. We will show that H quasi-isometrically embeds in $\mathcal{X}$. By Section 2.2 this implies that $MCG(\Sigma)$ quasi-isometrically embeds in a finite product of quasi-trees. The H-orbit of $\widehat{x}$ in $\mathcal{X}$ gives a Lipschitz embedding of H in $\mathcal{X}$. For the lower bound we have

$$\frac{1}{4} \sum_i \sum_{\gamma \in \mathcal{A}_i} d_\gamma(x_i, h(x_i))_K \leqslant d_{\mathcal{X}}(\widetilde{x}, h(\widetilde{x}))$$

for all $h \in H$ by the distance formula. By Proposition 4.18 we have

$$\sum_{Y \subseteq \Sigma} d_Y^T(\widetilde{x}, h(\widetilde{x})) \leqslant 2 \sum_i \sum_{\gamma \in \mathcal{A}_i} d_\gamma(x_i, h(x_i))_K + \frac{2T}{\xi} \sum_{Y \subseteq \Sigma} \left(d_Y(\widetilde{x}, \widehat{x})_\xi + d_Y(h(\widetilde{x}), h(\widehat{x}))_\xi \right).$$

The last term on the right is finite since $d_Y(\widetilde{x}, \widehat{x}) < \xi$ for all but finitely many Y and is independent of h since

$$d_Y(\widetilde{x}, \widehat{x})_\xi = d_{h(Y)}(h(\widetilde{x}), h(\widehat{x}))_\xi$$

by the group equivariance of the projections. The thick distance formula (Theorem 4.13) then gives a linear lower bound on the left hand side of the inequality in terms of the word length $|h|$. This completes the proof Theorem 1.2.

BIBLIOGRAPHY

[BBF15] Mladen Bestvina, Ken Bromberg, and Koji Fujiwara, *Constructing group actions on quasi-trees and applications to mapping class groups*, Publ. Math., Inst. Hautes Étud. Sci. **122** (2015), 1–64. ↑686, 688, 695, 696, 698, 707

[BBFS19] Mladen Bestvina, Ken Bromberg, Koji Fujiwara, and Alessandro Sisto, *Acylindrical actions on projection complexes*, Enseign. Math. **65** (2019), no. 1-2, 1–32. ↑686, 688

[BdlHV08] Bachir Bekka, Pierre de la Harpe, and Alain Valette, *Kazhdan's property (T)*, New Mathematical Monographs, vol. 11, Cambridge University Press, 2008. ↑686

[BDS07] Sergei Buyalo, Alexander Dranishnikov, and Viktor Schroeder, *Embedding of hyperbolic groups into products of binary trees*, Invent. Math. **169** (2007), no. 1, 153–192. ↑686

[BHS17] Jason Behrstock, Mark F. Hagen, and Alessandro Sisto, *Asymptotic dimension and small-cancellation for hierarchically hyperbolic spaces and groups*, Proc. Lond. Math. Soc. **114** (2017), no. 5, 890–926. ↑686

[CR07] Young-Eun Choi and Kasra Rafi, *Comparison between teichmüller and Lipschitz metrics*, J. Lond. Math. Soc. **76** (2007), no. 3, 739–756. ↑696

[DGO17] François Dahmani, Vincent Guirardel, and Denis V. Osin, *Hyperbolically embedded subgroups and rotating families in groups acting on hyperbolic spaces*, vol. 245, Memoirs of the American Mathematical Society, no. 1156, American Mathematical Society, 2017. ↑701

[DJ99] A. Dranishnikov and T. Januszkiewicz, *Every Coxeter group acts amenably on a compact space*, Proceedings of the 1999 Topology and Dynamics Conference (Salt Lake City, UT), vol. 24, 1999, pp. 135–141. ↑686

[FL08] Thomas Foertsch and Alexander Lytchak, *The de Rham decomposition theorem for metric spaces*, Geom. Funct. Anal. **18** (2008), no. 1, 120–143. ↑686

[Gro87] Mikhael L. Gromov, *Hyperbolic groups*, Essays in group theory, Mathematical Sciences Research Institute Publications, vol. 8, Springer, 1987, pp. 75–263. ↑693

[Gro93] M. Gromov, *Asymptotic invariants of infinite groups*, London Mathematical Society Lecture Note Series, vol. 182, Cambridge University Press, 1993. ↑686

[Gro75] Edna K. Grossman, *On the residual finiteness of certain mapping class groups*, J. Lond. Math. Soc. **9** (1974/75), 160–164. ↑702

[Hae20] Thomas Haettel, *Hyperbolic rigidity of higher rank lattices*, Ann. Sci. Éc. Norm. Supér. **53** (2020), no. 2, 439–468. ↑686

[Hum17] David Hume, *Embedding mapping class groups into a finite product of trees*, Groups Geom. Dyn. **11** (2017), no. 2, 613–647. ↑686

[LM07] Christopher J. Leininger and D. B. McReynolds, *Separable subgroups of mapping class groups*, Topology Appl. **154** (2007), no. 1, 1–10. ↑702

[MM99] Howard A. Masur and Yair N. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math. **138** (1999), no. 1, 103–149. ↑696

[MM00] _____, *Geometry of the complex of curves. II. Hierarchical structure*, Geom. Funct. Anal. **10** (2000), no. 4, 902–974. ↑696, 697

[NR97] Graham Niblo and Lawrence Reeves, *Groups acting on* CAT(0) *cube complexes*, Geom. Topol. **1** (1997), 1–7. ↑686

[RS11] Kasra Rafi and Saul Schleimer, *Curve complexes are rigid*, Duke Math. J. **158** (2011), no. 2, 225–246. ↑696

[Wat16] Yohsuke Watanabe, *Uniform local finiteness of the curve graph via subsurface projections*, Groups Geom. Dyn. **10** (2016), no. 4, 1265–1286. ↑696

Mladen BESTVINA
Department of Mathematics,
University of Utah,
Salt Lake City, UT 84112, (USA)

bestvina@math.utah.edu

Ken BROMBERG
Department of Mathematics,
University of Utah,
Salt Lake City, UT 84112, (USA)

bromberg@math.utah.edu

Koji FUJIWARA
Department of Mathematics,
Kyoto University,
Kyoto, 606-8502, (Japan)

kfujiwara@math.kyoto-u.ac.jp