# LINEARLY BOUNDED CONJUGATOR PROPERTY
# FOR MAPPING CLASS GROUPS

### Jing Tao

**Abstract.** Given two conjugate mapping classes $f$ and $g$, we produce a conjugating element $\omega$ such that $|\omega| \leq K(|f| + |g|)$, where $|\cdot|$ denotes the word metric with respect to a fixed generating set, and $K$ is a constant depending only on the generating set. As a consequence, the conjugacy problem for mapping class groups is exponentially bounded.

## 1 Introduction

Two fundamental problems in group theory posed by Dehn [Deh11] are the word problem and the conjugacy problem. Given a group with a fixed presentation, the *word problem* asks if there is an algorithm that can decide in finite time if a given word is the identity. The *conjugacy problem* seeks an algorithm to decide if two words represent the same conjugacy class. Since the conjugacy class of the identity element is itself, the word problem can be seen as a special case of the conjugacy problem. Not all groups have solvable word problem [Nov58, Boo59], hence the same is true for the conjugacy problem.

In this paper, we are interested in these problems for mapping class groups $\mathcal{MCG}(\mathcal{S})$ of surfaces $\mathcal{S}$ of finite type. We establish the following:

**Theorem A.** *There is an exponential-time algorithm to solve the conjugacy problem for $\mathcal{MCG}(\mathcal{S})$.*

There is some history to the word and conjugacy problems for $\mathcal{MCG}(\mathcal{S})$. The first solution to the word problem can be attributed to Grossman [Gro75], whose actual contribution is proving residual finiteness for $\mathcal{MCG}(\mathcal{S})$. In [Mos95], Mosher showed $\mathcal{MCG}(\mathcal{S})$ admits an automatic structure, from which a quadratic-time solution to the word problem is obtained. (See [ECH+92] for a background on automatic groups. It is not yet known if a sub-quadratic solution is possible.) In [Hem79], Hemion solved the conjugacy problem for $\mathcal{MCG}(\mathcal{S})$, but his algorithm is not exponentially bounded. In [Mos86], Mosher gave a faster algorithm for deciding conjugacy among pseudo-Anosov mapping classes. (A similar result was recently obtained by Agol [Ago11].) Using the work of Bestvina and Handel [BH95], which gives an algorithm for detecting pseudo-Anosov mapping classes, Mosher [Mos03] extended his result to compute complete conjugacy invariants for all mapping classes.

Our strategy to prove Theorem A is to apply Mosher's automaticity result. In general, a solution to the word problem does not necessarily yield a solution to the conjugacy problem: it is an open question whether all automatic groups have solvable conjugacy problem [ECH$^+$92]. A sufficient condition is if the group has *linearly bounded conjugator* (*L.B.C.*) *property* (see theorem below or Definition 2.2.1). The main theorem of our paper is that L.B.C. property is satisfied by $\mathcal{MCG}(\mathcal{S})$. This answers a question in [Far06].

**Theorem B** (L.B.C. property for $\mathcal{MCG}(\mathcal{S})$). *Let $\Lambda$ be a finite generating set for $\mathcal{MCG}(\mathcal{S})$. There exists a constant $K$, depending only on $\Lambda$, such that if $f, g \in \mathcal{MCG}(\mathcal{S})$ are conjugate, then there is a conjugating element $\omega$ with*

$$|\omega| \leq K(|f| + |g|).$$

To see how Theorem A follows from Theorem B, we give an algorithm to the conjugacy problem. Given two arbitrary elements $f, g \in \mathcal{MCG}(\mathcal{S})$, let $B$ be the ball of radius $K(|f| + |g|)$ in $\mathcal{MCG}(\mathcal{S})$. To decide if $f$ and $g$ are conjugate it suffices to check if $\omega \in B$ satisfies $\omega f \omega^{-1} g^{-1} = 1$. We run Mosher's quadratic algorithm to the word problem to all words of the form $\omega f \omega^{-1} g^{-1}$ with $\omega \in B$. The number of elements in $B$ is an exponential function of the radius, therefore the complexity of this solution is an exponential function of the word lengths of $f$ and $g$.

Linearly bounded conjugator property is satisfied by hyperbolic groups [Lys89, Lemma 10], as well as by torsion elements in groups acting on CAT(0) spaces [BH99, III.1.13]. These are important classes of groups that have solvable word and conjugacy problems [Gro87,BH99]. Hyperbolic groups in fact have efficient algorithms: the word problem is solvable in linear time, and the conjugacy problem in quadratic time [BH99]. As long as the surface $\mathcal{S}$ has disjoint isotopy classes of curves, $\mathcal{MCG}(\mathcal{S})$ is not hyperbolic, as Dehn twists about disjoint curves give rise to higher rank free abelian subgroups. It is also known that $\mathcal{MCG}(\mathcal{S})$ does not act on any complete CAT(0) space [BH99, II.7.26]. Nevertheless, $\mathcal{MCG}(\mathcal{S})$ shares many properties with hyperbolic groups, and much of the pursuit in its study has been to understand to what extent it resembles and differs from hyperbolic groups. Establishing L.B.C. property for $\mathcal{MCG}(\mathcal{S})$ thus provides another positive analogy between $\mathcal{MCG}(\mathcal{S})$ and hyperbolic groups.

After we announced our result, Hamenstädt [Ham09] announced biautomaticity for $\mathcal{MCG}(\mathcal{S})$, which generalizes Mosher's automaticity result as well as obtains Theorem A. Another consequence of her work is the exponentially bounded conjugator property for $\mathcal{MCG}(\mathcal{S})$. Notice, however, that this bound only gives a doubly exponential solution to the conjugacy problem if we use the same algorithm as described below Theorem B, since the search space for the conjugator would grow doubly exponential in terms of the word lengths of the elements.

**1.1  Idea of the proof of Theorem B.**  The proof of Theorem B is broken up into three arguments, following the classification of the elements of $\mathcal{MCG}(\mathcal{S})$

into pseudo-Anosov, reducible, and finite order. The case of the pseudo-Anosov elements was settled by Masur and Minsky [MM00], using the machinery of hierarchies developed in the same paper. This paper resolves the other two cases.

Surprisingly, it turns out the most delicate case involves the finite order elements of $\mathcal{MCG}(\mathcal{S})$. In many ways, pseudo-Anosov elements of $\mathcal{MCG}(\mathcal{S})$ can be viewed as the "hyperbolic" elements of $\mathcal{MCG}(\mathcal{S})$, whereas the finite order elements are the "elliptic" ones. The methods that Masur and Minsky developed are suited for elements that behave more hyperbolically, and thus are not effective for the finite order elements. Our main contribution is the development of new tools for the study of finite order mapping classes. Just as in the case of pseudo-Anosov mapping classes, we rely heavily on the machinery of hierarchies, which we need to extend so it is more suited to deal with the elliptic geometry.

We briefly explain how hierarchies are related to words in $\mathcal{MCG}(\mathcal{S})$. A natural model space for $\mathcal{MCG}(\mathcal{S})$ is the marking graph $\mathrm{Mark}(\mathcal{S})$ of $\mathcal{S}$. A marking $\mu_B \in \mathrm{Mark}(\mathcal{S})$ is a collection of curves on $\mathcal{S}$ satisfying certain technical conditions (see Section 2 for a precise definition). Given an element $f \in \mathcal{MCG}(\mathcal{S})$, the image of $\mu_B$ under $f$ determines $f$ up to finitely many choices. Being a model space, paths from $\mu_B$ to $f\mu_B$ in $\mathrm{Mark}(\mathcal{S})$ are naturally associated to words representing $f$, and the distance between $\mu_B$ and $f\mu_B$ is comparable to the word length of $f$ (fixing a generating set for $\mathcal{MCG}(\mathcal{S})$). Thus to understand the word length of $f$ is the same as understanding efficient paths from $\mu_B$ to $f\mu_B$.

Even though $\mathrm{Mark}(\mathcal{S})$ (or any other model space) is not hyperbolic or CAT(0), there is a coarsely well-defined projection map from $\mathrm{Mark}(\mathcal{S})$ to a product of hyperbolic spaces $\prod_Z \mathcal{C}(Z)$: each factor $\mathcal{C}(Z)$ is the curve complex of a subsurface $Z$ of $\mathcal{S}$, and the product is taken over all essential (possibly annular) subsurfaces of $\mathcal{S}$. The fact that each $\mathcal{C}(Z)$ is hyperbolic was established by Masur and Minsky [MM99]. For each such $Z$, the projection map $\pi_Z \colon \mathrm{Mark}(\mathcal{S}) \to \mathcal{C}(Z)$ is obtained by a surgery procedure (see Section 2). By connecting $\pi_Z(\mu_B)$ and $\pi_Z(f\mu_B)$ by a geodesic path in $\mathcal{C}(Z)$, one can associate to each element $f \in \mathcal{MCG}(\mathcal{S})$ a family of geodesics in curve complexes. These geodesics are organized by hierarchies to produce efficient paths connecting $\mu_B$ to $f\mu_B$ in $\mathrm{Mark}(\mathcal{S})$. An important consequence of hierarchies is the distance formula (Theorem 2.6.5), which states that the distance between $\mu_B$ and $f\mu_B$ is well approximated by the sum of the curve complex distances between their projections, where the sum is taken over those subsurfaces to which the projections are sufficiently far apart.

The hyperbolic geometry of the pseudo-Anosov elements of $\mathcal{MCG}(\mathcal{S})$ is exhibited in the fact that they act hyperbolically (with north–south dynamics) on the curve complex $\mathcal{C}(\mathcal{S})$ of $\mathcal{S}$ [MM99, MM00]. This is analogous to the way how the infinite order elements of a hyperbolic group act on a Cayley graph of the group. Hierarchies were used to build quasi-axes in $\mathcal{C}(\mathcal{S})$ for the action of pseudo-Anosov mapping classes. There is also a fellow-traveling type property for hierarchies that applies to fellow-traveling quasi-axes. These facts allowed Masur and Minsky

to extend the proof of L.B.C. property from infinite-order elements of hyperbolic groups to pseudo-Anosov elements of $\mathcal{MCG}(\mathcal{S})$ [MM00, Theorem 7.2].

The appropriate analogy for the finite order elements of $\mathcal{MCG}(\mathcal{S})$ are the torsion elements of a group $G$ that acts properly and cocompactly on a CAT(0) (or hyperbolic) space $X$. We are inspired by the argument contained in [BH99] on L.B.C. property for torsion elements of $G$ which we will briefly sketch. Let $x \in X$ be a fixed base point. We say an element $g \in G$ acts *elliptically* on $X$ if it satisfies two conditions. First, $g$ acts on $X$ with (coarse) fixed points. Second, the distance from $x$ to the center of mass of the orbit of $x$ under $\langle g \rangle$ is comparable to the word length of $g$. When $X$ is CAT(0) (or hyperbolic), $g$ is torsion implies $g$ acts elliptically. After conjugating $g$ by an appropriate element, its center of mass can be moved into a fixed ball containing a fundamental domain for the action of $G$ on $X$. The set of torsion elements of $G$ having fixed points inside the ball is finite and contains a representative for each conjugacy class. From here, one can reduce L.B.C. property in the elliptic case to a finite set.

To establish L.B.C. property for finite order elements of $\mathcal{MCG}(\mathcal{S})$, we also show they act elliptically on $\mathrm{Mark}(\mathcal{S})$. More precisely,

**Theorem C.** *Let $\mu_B \in \mathcal{MCG}(\mathcal{S})$ be a fixed base point. There exist constants $R$ and $k$ depending only on $\mu_B$ such that the following holds. For any finite order element $f \in \mathcal{MCG}(\mathcal{S})$, there exists $\mu \in \mathrm{Mark}(\mathcal{S})$ such that $\mu$ is an $R$-fixed point of $f$ (i.e. $d_{\mathrm{Mark}(\mathcal{S})}(\mu, f\mu) \leq R$) and*

$$d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \mu) \leq k|f|. \tag{1}$$

COROLLARY D. *There exist a constant $K$, depending only on $\mathcal{S}$, and a finite set of elements $\Sigma \subset \mathcal{MCG}(\mathcal{S})$ such that if $f \in \mathcal{MCG}(\mathcal{S})$ has finite order, then there exists $\omega \in \mathcal{MCG}(\mathcal{S})$ such that $\omega f \omega^{-1} \in \Sigma$ and $|\omega| \leq K|f|$.*

The proof of Theorem C is the technical part of this paper. It is easy to see that there exists a constant $R_1$, depending only on $\mathcal{S}$, such that any finite order element $f \in \mathcal{MCG}(\mathcal{S})$ acts on $\mathrm{Mark}(\mathcal{S})$ with $R_1$-fixed points. The hard part is finding an $R_1$-fixed point of $f$ that lies "$k$-close" (in the sense of (1)) to $\mu_B$, for some uniform $k$. Using the projection maps from $\mathrm{Mark}(\mathcal{S})$ to curve complexes, what we want is to find a marking $\mu \in \mathrm{Mark}(\mathcal{S})$ such that, for any $Z, \pi_Z(\mu)$ lies sufficiently close to the convex hull of $\{\pi_Z(f^i\mu_B)\}$ (the projection to $\mathcal{C}(Z)$ of the orbit of $\mu_B$ under $\langle f \rangle$). To find such $\mu$, our strategy is to take an arbitrary $R_1$-fixed point $\mu'$ of $f$ and construct from it a marking $\mu$ (possibly equal to $\mu'$) that satisfies Theorem C for appropriate constants $k$ and $R$ ($R$ possibly bigger than $R_1$). The construction of $\mu$ is through a sequence of modifications on $\mu'$, taken place in subsurfaces of $\mathcal{S}$ to which the projections of $\mu'$ are "far" from the convex hull. (If in every subsurface of $\mathcal{S}$, the projection of $\mu$ is not "far" from the convex hull, then $\mu = \mu'$.) An important part of our proof that makes this process work are two technical lemmas (Section 3), which show that the symmetries of the action of $f$ on $\mathcal{S}$ can be detected by hierarchies.

To establish L.B.C. property for reducible elements of $\mathcal{MCG}(\mathcal{S})$, we combine the two arguments, for pseudo-Anosov elements and for finite order elements. If $f \in \mathcal{MCG}(\mathcal{S})$ is a reducible element of infinite order, then up to taking powers the surface $\mathcal{S}$ can be decomposed into a collection of subsurfaces on which $f$ is either pseudo-Anosov or has finite order. In order to apply induction to subsurfaces, we need to built paths from $\mu_B$ to $f\mu_B$ in $\mathrm{Mark}(\mathcal{S})$ that move only in the complementary subsurfaces of the reducing system of $f$. This is possible if the initial marking $\mu_B$ contains the reducing system of $f$. However, one marking cannot contain all possible reducing systems, even up to conjugation. But it suffices to reduce to a finite problem. We show:

**Theorem E.**   *There exist a constant $k$ and a finite set of markings $\mathcal{M}$ so that if $f \in \mathcal{MCG}(\mathcal{S})$ is reducible, then there exists $\omega \in \mathcal{MCG}(\mathcal{S})$ such that the reducing system of $\omega f \omega^{-1}$ is contained in some $\mu \in \mathcal{M}$ and $|\omega| \leq k|f|$.*

Finally, each case in the classification will produce a different constant. The proof of L.B.C. property for $\mathcal{MCG}(\mathcal{S})$ will be completed by taking a maximum over the three constants.

The organization of the paper is as follows.

- In Section 2, we review basic definitions and the theory of hierarchies. A key notion that we will introduce is the notion of *separating markings* in Section 2.9.
- In Section 3, we give a couple of definitions and prove two technical lemmas about finite order mapping classes that will be useful for the next section. We also construct an example that motivates this section and the next section.
- In Section 4, we prove Theorem C and derive L.B.C. property for finite order mapping classes.
- In Section 5, we prove Theorem E and use the known results for pseudo-Anosov and finite order elements to derive L.B.C. property for infinite order reducible mapping classes.

## 2   Preliminaries

In this section, we develop the background material for the paper. Our main tool will be Masur and Minsky's theory of hierarchies. From Sections 2.6 to 2.8, we will summarize the properties of hierarchies that will be needed for this paper. Some of the definitions will be merely sketched and most of the proofs will be omitted. We refer the reader to Masur–Minsky's paper [MM00] for more details. We also refer to [FLP79, FM12] for general references on mapping class groups and the topology of surfaces, and to [Gro87, BH99] for references on $\delta$-hyperbolic spaces.

**2.1   Arcs, curves, surfaces and subsurfaces.**   Let $\mathcal{S} = \mathcal{S}_{g,p}$ be a connected, oriented surface of genus $g$ with $p$ punctures. We call $\xi(\mathcal{S}) = 3g - 3 + p$ the *complexity* of $\mathcal{S}$. Surfaces of complexity strictly greater than 1 are called *generic* surfaces. Surfaces of complexity 1 are called *sporadic* and they are topologically either

the four-holed sphere or one-holed torus. Two remaining low-complexity cases are
*exceptional surfaces*. Complexity 0 is the three-holed sphere or a *pair of pants*, and
complexity $-1$ is topologically an annulus.

Throughout this paper we will be working with a generic surface $\mathcal{S}$ without
boundary. But sporadic and exceptional surfaces and surfaces with boundary natu-
rally arise as subsurfaces of $\mathcal{S}$, and thus are important for induction arguments.

An *essential curve* or just a *curve* on $\mathcal{S}$ will always mean the free isotopy class
of a simple closed curve that is not null-homotopic or homotopic to a puncture or a
boundary component. A *multicurve* or a *curve system* will mean a finite collection
of distinct curves that can be realized disjointly. A *pants decomposition* of $\mathcal{S}$ is a
maximal curve system $c$ on $\mathcal{S}$. In particular, each component of $\mathcal{S} \setminus c$ is topologically
a pair of pants. Note that a pants decomposition exists for $\mathcal{S}$ if and only if $\xi(\mathcal{S}) \geq 1$,
in which case the cardinality (or the number of curves) of $c$ is equal to $\xi(\mathcal{S})$.

To talk about arcs we need $\mathcal{S}$ to have boundary. An *arc* on $\mathcal{S}$ will be an isotopy
class of a simple arc $\delta$, with isotopies relative to the boundary, such that $\delta$ has both
endpoints on $\partial \mathcal{S}$ and is not isotopic to a boundary component.

The (*geometric*) *intersection number* $i(\alpha, \beta)$ of a pair of curves $\alpha$ and $\beta$ will be the
minimal number of intersections among representatives of $\alpha$ and $\beta$. The geometric
intersection number between two arcs on $\mathcal{S}$ will be the minimal number of inter-
sections in the interior of $\mathcal{S}$ modulo isotopies relative to $\partial \mathcal{S}$. Note that intersection
number of an arc or curve with itself is always zero.

A *subsurface* $Y$ of $\mathcal{S}$ is the isotopy class of a closed and connected subsurface of
$\mathcal{S}$ that is incompressible and non-peripheral. We include the possibility that $Y = \mathcal{S}$
unless we say a *proper* subsurface. By $\partial Y$ we will mean the multicurve comprised of
the boundary components of a representative of $Y$. An annular subsurface $A$ of $\mathcal{S}$
is a regular neighborhood of a curve $\alpha$ with simple boundaries. We will often abuse
terminology by confusing $A$ with its *core curve* $\alpha$, and refer to $\alpha$ as a subsurface of
$\mathcal{S}$ as well. In this case, $\partial A$ will mean $\alpha$. For reasons we shall see, we will distinguish
subsurfaces that are not pants, called *essential subsurfaces* or *domains*.

Given a curve $\alpha$ and a non-annular domain $Y$ of $\mathcal{S}$, we will say $\alpha$ is disjoint from
$Y$ if it can be homotoped away from a representative of $Y$. Note that this includes
the case that $\alpha$ is a curve in $\partial Y$. If $\alpha$ can be realized as an essential curve in a
representative of $Y$, then we will say $\alpha$ is a curve in $Y$. In all other cases, we will
say $\alpha$ *crosses* $Y$. For an annular domain $A$ with core curve $\beta$, then we have the
possibilities that $\alpha$ is disjoint from $A$ if $\alpha$ and $\beta$ are disjoint, or $\alpha$ crosses $A$ if $\alpha$ and
$\beta$ intersect.

Similarly, given two domains $Y$ and $Z$ of $\mathcal{S}$, we will say $Y$ and $Z$ are: *disjoint* if
$Y$ and $Z$ can be homotoped to be disjoint from each other; *nested* if $Y$ and $Z$ can be
homotoped so that either $Y$ is contained in $Z$ or $Z$ is contained in $Y$; and *interlock*
if they are neither disjoint or nested. Note that when $Y = A$ is an annular domain
with core curve $\alpha$, then $A$ and $Z$ being disjoint is consistent with $\alpha$ and $Z$ being
disjoint, $A$ and $Z$ are nested if $\alpha$ is contained $Z$ (and $Z$ is not annular), and, finally,
$A$ and $Z$ interlock if $\alpha$ crosses $Z$.

**2.2 Mapping class groups.** Let $\mathrm{Homeo}^+(\mathcal{S})$ be the group of orientation-preserving self-homeomorphisms of $\mathcal{S}$. The *mapping class group* of $\mathcal{S}$ is

$$\mathcal{MCG}(\mathcal{S}) = \mathrm{Homeo}^+(\mathcal{S})/\sim$$

where $f \sim g$ if and only if $g^{-1} \circ f$ is isotopic to the identity map on $\mathcal{S}$. Elements of $\mathcal{MCG}(\mathcal{S})$ are called *mapping classes*. It is well-known that $\mathcal{MCG}(\mathcal{S})$ is finitely generated (and finitely presented) [Lic64]. For this paper, we will fix a finite generating set $\Lambda$ of $\mathcal{MCG}(\mathcal{S})$. We will often regard $\mathcal{MCG}(\mathcal{S})$ as a metric space by considering the word metric $|\cdot| = |\cdot|_\lambda$ induced by $\Lambda$.

If $\mathcal{S}$ is a once-punctured torus or four-times punctured sphere, then $\mathcal{MCG}(\mathcal{S})$ is commensurable to $\mathrm{SL}(2, \mathbb{Z})$. The mapping class group of a thrice-punctured sphere is finite. For us, an annulus $A$ will always appear as a regular neighborhood of a simple closed curve on an ambient surface, so $A$ has two boundary components. Let $\mathcal{MCG}(A, \partial A)$ be the group of isotopy classes of homeomorphisms of $A$ relative to $\partial A$. One checks that $\mathcal{MCG}(A, \partial A)$ is homeomorphic to $\mathbb{Z}$.

DEFINITION 2.2.1 (L.B.C. property). *Given a finitely generated group $G$ equipped with a finite generating set $\Lambda$, we say a conjugacy class $\mathfrak{c}$ of $G$ has linearly bounded conjugators if for any $f, g \in \mathfrak{c}$, there exists a conjugating element $\omega \in G$ such that*

$$|\omega| \leq K_\mathfrak{c}(|f| + |g|),$$

*where $|\cdot|$ represent the word length in $\Lambda$, and $K_\mathfrak{c}$ depends only on $\mathfrak{c}$ and $\Lambda$. If $K = K_\mathfrak{c}$ can be taken to be independent of the conjugacy class $\mathfrak{c}$, then we say $G$ has linearly bounded conjugator property or L.B.C. property. If $G$ has L.B.C. property for $\Lambda$, then changing $\Lambda$ to any other finite generating set changes $K$ by a bounded amount. Therefore, this definition is independent of the choice of the generating set, so $\Lambda$ can always be taken to be a symmetric generating set.*

Mapping class groups of non-generic surfaces satisfy L.B.C. property. We would like to show the same is true for mapping class groups of generic surfaces. The first observation is that the Nielsen–Thurston classification of mapping classes is a conjugacy invariant. This means that we can argue for L.B.C. property separately for each type. We refer to [Thu88] and [FM12, §13] for more details on the classification theorem. Recall that a mapping class $f$ is called *irreducible* if $f$ does not fix any multicurve (setwise); otherwise $f$ is called *reducible*. The following statement applies to all surfaces $\mathcal{S}$.

**Theorem 2.2.2** (Nielsen–Thurston classification for $\mathcal{MCG}(\mathcal{S})$). *Every element $f \in \mathcal{MCG}(\mathcal{S})$ is either pseudo-Anosov, periodic (finite order), or reducible. Furthermore, for each $f \in \mathcal{MCG}(\mathcal{S})$, there exists a (possibly empty) multicurve $\sigma$ invariant under $f$ with the following property. Let $Y_1, \ldots, Y_k$ be the connected components of $\mathcal{S} \setminus \sigma$, and, for each $i$, choose the smallest $n_i \in \mathbb{N}$ so that $f^{n_i}(Y_i) = Y_i$. Then for any $i$, $f^{n_i}|_{Y_i}$ either has finite order or is pseudo-Anosov.*

The multicurve $\sigma$ satisfying Theorem 2.2.2 $f$ is called a *reducing system* for $f$. For each $Y_i \in \mathcal{S} \setminus \sigma$, the map $f^{n_i}$ is called the *first return map of $f$ to $Y_i$*. Note that the first return map of $f$ to $Y_i$ means exactly that $f^{n_i}|_{Y_i}$ can be viewed as an element of $\mathcal{MCG}(Y_i)$. The content of the classification can be rephrased to say $f$ is pseudo-Anosov if and only if $f$ is irreducible of infinite order.

DEFINITION 2.2.3 (Canonical reducing systems). *By choosing $\sigma$ to be a minimal collection of curves satisfying Theorem 2.2.2, then $\sigma = \sigma_f$ is unique up to isotopy and is called the canonical reducing system for $f$. If $f \in \mathcal{MCG}(\mathcal{S})$ is either pseudo-Anosov or finite order, then $\sigma_f = \emptyset$ (see [Mos07]).*

In [MM00, §7], Masur–Minsky established L.B.C. property for the pseudo-Anosov elements of $\mathcal{MCG}(\mathcal{S})$.

**Theorem 2.2.4** (L.B.C. property for pseudo-Anosov mapping classes). *There exists a constant $K$, depending only on $\mathcal{S}$, such that if $f, g \in \mathcal{MCG}(\mathcal{S})$ are conjugate pseudo-Anosov mapping classes, then there is a conjugating element $\omega \in \mathcal{MCG}(\mathcal{S})$ with*

$$|\omega| \leq K(|f| + |g|).$$

Our goal in this paper is to prove L.B.C. property for the finite order and reducible elements of $\mathcal{MCG}(\mathcal{S})$. The argument for finite order mapping classes is the hard part of this paper. The argument for reducible mapping classes is inductive and will make use of the canonical reducing system.

**2.3  Complexes of curves.**  The *complex of curves* $\mathcal{C}(\mathcal{S})$ on a surface $\mathcal{S}$ is a locally infinite, finite dimensional simplicial complex on which $\mathcal{MCG}(\mathcal{S})$ acts by automorphisms. Its definition first appeared in [Har81]. We treat generic, sporadic, and exceptional surfaces separately.

- **Generic surfaces** Suppose $\mathcal{S}$ has $\xi(\mathcal{S}) > 1$. The $k$th skeleton $\mathcal{C}_k(\mathcal{S})$ consists of all curve systems on $\mathcal{S}$ of cardinality $k + 1$. There is an obvious inclusion of $\mathcal{C}_{k-1}(\mathcal{S}) \hookrightarrow \mathcal{C}_k(s)$ by face relations. Top dimensional simplices of $\mathcal{C}(\mathcal{S})$ correspond to pants decompositions on $\mathcal{S}$, hence $\dim(\mathcal{C}(\mathcal{S})) = \xi(\mathcal{S}) - 1$.
- **Sporadic surfaces** With the above definition, the curve complex of a sporadic surface $\mathcal{S}$ would be a disconnected set of points. To construct a more useful object, we modify the definition to allow two vertices in $\mathcal{C}(\mathcal{S})$ span an edge if they intersect minimally over $\mathcal{S}$ (once for one-holed torus and twice for four-holed sphere). It is a classical theorem that with this definition $\mathcal{C}(\mathcal{S})$ is isomorphic to the Farey graph [HT80, Min96].
- **Pants** A pair of pants has no essential curves. Here we do not modify the definition and let the curve complex of pants be empty. This is the reason why we do not consider pants to be essential subsurfaces.
- **Annuli** An arbitrary annulus has no essential curves. But for us, an annulus $A$ will always appear as a regular neighborhood of a curve $\gamma$ in a larger surface $\mathcal{S}$,

and we would like $\mathcal{C}(A)$ (or $\mathcal{C}(\gamma)$) to record twist information about $\gamma$. Vertices of $\mathcal{C}(A)$ will be properly embedded arcs and two arcs are connected by an edge if they can be isotoped relative endpoints to have disjoint interiors.

By an element or subset of $\mathcal{C}(\mathcal{S})$ we will always mean an element or subset of $\mathcal{C}_0(\mathcal{S})$. We make $\mathcal{C}(\mathcal{S})$ into a complete geodesic metric space by endowing each simplex with an Euclidean structure with edge lengths 1. From the perspective of coarse geometry, we do not lose anything by identifying $\mathcal{C}(\mathcal{S})$ with its 1-skeleton. We denote by $d_{\mathcal{C}(\mathcal{S})}$, or more simply by $d_{\mathcal{S}}$, the shortest distance in $\mathcal{C}_1(\mathcal{S})$ between two vertices. If $A$ is an annulus with a core curve $\gamma$, we will also use the notation $d_\gamma$ or $d_A$ to denote distances in $\mathcal{C}(A)$. For any surface $\mathcal{S}$ including annuli, induction on intersection number can be used to show $\mathcal{C}(\mathcal{S})$ is connected, and $d_{\mathcal{S}}(\alpha,\beta) \leq 2i(\alpha,\beta)+1$ (see [MM99, MM00]). The simplicial action of $\mathcal{MCG}(\mathcal{S})$ on $\mathcal{C}(\mathcal{S})$ preserves this metric. The action is not proper. The quotient $\mathcal{C}(\mathcal{S})/\mathcal{MCG}(\mathcal{S})$ parametrizes curves on $\mathcal{S}$ up to homeomorphisms, hence it is finite.

For a generic surface $\mathcal{S}$, $d_{\mathcal{S}}$ coarsely measures the complexity between two curves in the following sense: $d_{\mathcal{S}}(\alpha,\beta) = 1$ if and only if $\alpha$ and $\beta$ are disjoint; $d_{\mathcal{S}}(\alpha,\beta) = 2$ if and only if $\alpha$ and $\beta$ cohabit a proper subsurface $Y \subset \mathcal{S}$; $d_{\mathcal{S}}(\alpha,\beta) \geq 3$ if and only $\alpha$ and $\beta$ *fill* $\mathcal{S}$, or the complement of their union in $\mathcal{S}$ does not support any essential curve.

The following theorem in [MM99] gives us some geometric control over paths in $\mathcal{C}(\mathcal{S})$.

**Theorem 2.3.1** [MM99]. *For any surface $\mathcal{S}$ that is not a pair of pants, $\mathcal{C}(\mathcal{S})$ has infinite diameter and is $\delta$-hyperbolic.*

For sporadic surfaces, Theorem 2.3.1 follows from a classical result that the Farey graph is quasi-isometric to an infinite-valence tree (see [Man05]). In the case of an annulus $A$, Theorem 2.3.1 follows from the fact that $\mathcal{C}(A)$ is quasi-isometric to $\mathbb{Z}$ (see [MM00, §2.4]).

For generic surfaces, there are several ways to see that $\mathcal{C}(\mathcal{S})$ has infinite diameter. Relevant to our paper is the following lemma.

LEMMA 2.3.2 [MM99, Proposition 4.6]. *There exists $k = k(\mathcal{S})$ such that for any pseudo-Anosov $f \in \mathcal{MCG}(\mathcal{S})$, any vertex $v \in \mathcal{C}(\mathcal{S})$, and any $n \in \mathbb{Z}$,*

$$d_{\mathcal{S}}(v, f^n(v)) \geq k|n|.$$

The proof of $\delta$-hyperbolicity of $\mathcal{C}(\mathcal{S})$ for a generic $\mathcal{S}$ is nontrivial. We also refer to [Bow06] for an alternate proof.

**2.4 Subsurface projections.** In this section, we restrict our discussion to domains of an ambient surface $\mathcal{S}$ with $\xi(\mathcal{S}) \geq 1$. To do away with isotopy classes of curves and surfaces, we will equip $\mathcal{S}$ with hyperbolic metric so that we may consider geodesic representatives for curves and (non-annular) subsurfaces of $\mathcal{S}$ bounded by them.

Let $Y \subset \mathcal{S}$ be a proper domain. There is a map

$$\pi_Y : \mathcal{C}(\mathcal{S}) \to \mathcal{P}(\mathcal{C}(Y)),$$

taking an element of $\mathcal{C}(\mathcal{S})$ to a subset of $\mathcal{C}(Y)$ of bounded diameter. We call $\pi_Y(\alpha)$ the *projection* of $\alpha$ to $Y$. Note that in the definition below, the projection map also makes sense if we replace $\mathcal{S}$ by any subsurface of $\mathcal{S}$ that contains $Y$ as a proper subsurface.

We first define the projection to a non-annular domain $Y$. If $\alpha$ and $Y$ are disjoint, then $\pi_Y(\alpha) = \emptyset$. If $\alpha$ is a curve in $Y$, then $\pi_Y(\alpha) = \{\alpha\}$. Otherwise, $\alpha$ crosses $Y$ and $\alpha \cap Y$ consist of a collection of arcs in $Y$. The endpoints of each arc $\delta \subset \alpha \cap Y$ lie on one or two components of $\partial Y$. Let $N$ be a regular neighborhood of the union of $\delta$ with its corresponding component(s) in $\partial Y$. $N$ has either one or two components that are essential in $Y$. Let $\pi_Y(\delta)$ be the set of boundary component(s) of $N$. We define

$$\pi_Y(\alpha) = \bigcup_{\delta \subset \alpha \cap Y} \pi_Y(\delta).$$

Now suppose $Y = A$ is an annulus with core curve $\gamma$. There is a unique annular cover of $\mathcal{S}$

$$p : \widehat{A} \to \mathcal{S}$$

to which $A$ lifts homeomorphically. Since $\mathcal{S}$ admits a hyperbolic metric, this cover has a natural compactification, also denote by $\widehat{A}$. We define $\mathcal{C}(A) = \mathcal{C}(\widehat{A})$. For any curve $\alpha$ in $\mathcal{S}$, components of $p^{-1}(\alpha)$ that are essential arcs form a subset in $\mathcal{C}(A)$. We will let $\pi_A(\alpha)$ be this corresponding set in $\mathcal{C}(A)$.

Denote by $\mathrm{diam}_Y(\cdot)$ the diameter of subsets in $\mathcal{C}(Y)$. For any two subsets $A, B \subset \mathcal{C}(Y)$, let

$$d_Y(A, B) = \mathrm{diam}_Y(A \cup B).$$

Given a pair of curves $\alpha, \beta \in \mathcal{C}(\mathcal{S})$ and a domain $Y \subset \mathcal{S}$, we define

$$d_Y(\alpha, \beta) = d_Y(\pi_Y(\alpha), \pi_Y(\beta)).$$

For any multicurve $\sigma$, one can also project $\sigma$ to $\mathcal{C}(Y)$ in the obvious way: $\pi_Y(\sigma) = \bigcup_{\alpha \in \sigma} \pi_Y(\alpha)$. Given two multicurves $\sigma$ and $\tau$, the distance $d_Y(\sigma, \tau)$ is similarly defined.

The follow result asserts that subsurface projections are coarsely well-defined and Lipschitz.

LEMMA 2.4.1 [MM00, Lemma 2.3]. *For any multicurve $\sigma$ on $\mathcal{S}$ and any domain $Y \subset \mathcal{S}$, if $\pi_Y(\sigma) \neq \emptyset$, then $\mathrm{diam}_Y(\pi_Y(\sigma)) \leq 2$.*

Suppose $Y$ and $Z$ are domains of $\mathcal{S}$ such that $Y$ is contained in $Z$. Then the maps $\pi_Y$ and $\pi_Y \circ \pi_Z$ are "coarsely equal" as maps from $\mathcal{C}(\mathcal{S}) \to \mathcal{P}(\mathcal{C}(Y))$.

LEMMA 2.4.2 [BKMM12, Lemma 2.12]. *There exists a constant $M$ depending only on $\mathcal{S}$ such that for any multicurve $\sigma$,*

$$\mathrm{diam}_Y(\pi_Y(\sigma), \pi_Y \circ \pi_Z(\sigma)) \leq M.$$

We also have the following contraction property for the projection map from [MM00, Theorem 3.1].

**Theorem 2.4.3** (Bounded geodesic image). *There exists a constant $M_0$ depending only on $\mathcal{S}$ such that the following holds. Suppose $Y \subset \mathcal{S}$ is a proper essential subsurface, and $g$ is a geodesic in $\mathcal{C}(\mathcal{S})$ such that $\pi_Y(v) \neq \emptyset$ for every vertex $v \in g$. Then*

$$\mathrm{diam}_Y(g) \leq M_0.$$

We say $g$ *cuts* $Y$ if $\pi_Y(v) \neq \emptyset$ for every vertex $v \in g$, and $g$ *misses* $Y$ otherwise. If $d_{\mathcal{S}}(g, \partial Y) \geq 2$ then $g$ cuts $Y$. On the other hand, by Theorem 2.4.3, if $u, v \in \mathcal{C}(\mathcal{S})$ has $d_Y(u, v) > M_0$, then any geodesic $g$ in $\mathcal{C}(\mathcal{S})$ between $u$ and $v$ misses $Y$.

**2.5 Marking graph.** Another useful combinatorial object that admits an action by $\mathcal{MCG}(\mathcal{S})$ is the *marking graph* $\mathrm{Mark}(\mathcal{S})$ of $\mathcal{S}$. Roughly, a *marking* $\mu$ on $\mathcal{S}$ is a multicurve $c$ on $\mathcal{S}$ with additionally a set of transverse curves that serve to record twisting data about each curve in $c$. Below, we give a precise definition that works for any surface $\mathcal{S}$ with $\xi(\mathcal{S}) \geq 1$.

A marking $\mu$ on $\mathcal{S}$ is a set of ordered pairs $\{(\alpha_i, t_i)\}$, where the *base curves* $\mathrm{base}(\mu) = \{\alpha_i\}$ is a multicurve on $\mathcal{S}$, and each *transversal* $t_i$ is either empty or is a diameter-1 set of vertices in $\mathcal{C}(\alpha_i)$. The set of transversals $\{t_i\}$ is denoted by $\mathrm{trans}(\mu)$. A transversal $t$ in the pair $(\alpha, t)$ is called *clean* if $t = \pi_\alpha(\beta)$, where $\beta$ is a curve on $\mathcal{S}$ such that $\alpha$ and $\beta$ are Farey-neighbors in the subsurface that they fill. A marking $\mu$ is *clean* if every non-empty transversal $t$ is clean, and the curve $\beta$ inducing $t$ does not intersect any other base curve other than $\alpha$. A marking $\mu$ is called *complete* if $\mathrm{base}(\mu)$ is a pants decomposition of $\mathcal{S}$ and no transversal is empty. If $\mu$ is complete and clean, then a transversal $t$ determines uniquely the curve $\beta$ such that $t = \pi_\alpha(\beta)$. If $\mu$ is not clean then there is bounded number ways of picking a *compatible* clean marking $\mu'$, in the following sense:

LEMMA 2.5.1 [MM00, Lemma 2.4]. *There exists a constant $M$ depending only on $\mathcal{S}$ satisfying the following. For any complete marking $\mu$ on $\mathcal{S}$, there exists a uniformly bounded number (depending only on $\mathcal{S}$) of complete clean markings $\mu'$ such that $\mathrm{base}(\mu) = \mathrm{base}(\mu')$, and $d_\alpha(t, t') \leq M$ for any $(\alpha, t) \in \mu$ and $(\alpha, t') \in \mu'$.*

We will often suppress the pair notation and regard a marking $\mu$ as the union of its base curves and transversals, i.e. $\mu = (\cup_{\alpha \in \mathrm{base}(\mu)} \alpha) \bigcup (\cup_{t \in \mathrm{trans}(\mu)} t)$.

DEFINITION 2.5.2 (Marking graph). *The marking graph $\mathrm{Mark}(\mathcal{S})$ is the graph with vertices representing complete clean markings on $\mathcal{S}$. Two vertices $\mu = \{(\alpha_i, \pi_{\alpha_i}(\beta_i))\}$ and $\mu' = \{(\alpha'_i, \pi_{\alpha'}(\beta'_i))\}$ are connected by an edge if they differ by one of the following elementary moves:*

- *Twist for some $i$, $\beta'_i$ is obtained from $\beta$ by a twist or half-twist along $\alpha_i$. All base curves and other transversals of $\mu$ and $\mu'$ agree.*
- *Flip let $\mu''$ be the (unclean) marking obtained from $\mu$ by "flipping" $(\alpha_i, \pi_{\alpha_i}(\beta_i))$ to $(\beta_i, \pi_{\beta_i}(\alpha_i))$, for some $i$. The marking $\mu'$ is any clean marking compatible with $\mu''$ replacing all transversals $\beta_j$ that intersect $\beta_i$.*

We equip $\mathrm{Mark}(\mathcal{S})$ with the combinatorial edge metric, denoted by $d_{\mathrm{Mark}(\mathcal{S})}$. Like $\mathcal{C}(\mathcal{S})$, $\mathrm{Mark}(\mathcal{S})$ is connected and admits an action of $\mathcal{MCG}(\mathcal{S})$ by isometries. But unlike $\mathcal{C}(\mathcal{S})$, $\mathrm{Mark}(\mathcal{S})$ is locally finite and the action of $\mathcal{MCG}(\mathcal{S})$ is proper. The quotient $\mathrm{Mark}(\mathcal{S})/\mathcal{MCG}(\mathcal{S})$ is also finite, since there are only finitely many complete clean markings up to homeomorphisms of $\mathcal{S}$ [MM00]. By a standard application of Švarc–Milnor, the orbit map $\mathcal{MCG}(\mathcal{S}) \to \mathrm{Mark}(\mathcal{S})$ is a quasi-isometry.

DEFINITION 2.5.3 (Projection of markings). *Let $Y \subset \mathcal{S}$ be essential and let $\mu \in \mathrm{Mark}(\mathcal{S})$. We can project $\mu$ to $Y$, also denoted by $\pi_Y(\mu)$, in the following way. Namely, if $Y$ is not a curve in $\mathrm{base}(\mu)$, then $\pi_Y(\mu) = \pi_Y(\mathrm{base}(\mu))$. If $Y = \alpha$ is a curve contained in $\mathrm{base}(\mu)$, then $\pi_Y(\mu) = t$, where $t$ is the transversal curve to $\alpha$ in $\mu$. Note that, since $\mu$ is a complete marking, the projection map is always non-empty.*

Since $\mathrm{base}(\mu)$ is a diameter-1 set in $\mathcal{C}(\mathcal{S})$, in light of Lemma 2.4.1 the projection map is Lipschitz:

LEMMA 2.5.4 [MM00, Lemma 2.5]. *For any $\mu, \nu \in \mathrm{Mark}(\mathcal{S})$ and any domain $Y \subseteq \mathcal{S}$,*

$$d_Y(\mu, \nu) \leq 4 \, d_{\mathrm{Mark}(\mathcal{S})}(\mu, \nu).$$

If $c$ is a multicurve and $\mu$ a marking with $c \subseteq \mathrm{base}(\mu)$, then we say $\mu$ is an *extension* of $c$. We will often start with a multicurve $c$ and *extend* it to a marking $\mu$. This amounts to choosing a marking on all the essential non-annular components of $\mathcal{S} \setminus c$ and choosing a transversal for each curve $\alpha \in c$. There are many ways to extend a marking in general, but most often we will need the marking $\mu$ to satisfy certain desired properties so those choices will be bounded.

DEFINITION 2.5.5 (Induced marking). *Let $Y \subset \mathcal{S}$ be an non-annular domain. We define a map*

$$\Pi_Y : \mathrm{Mark}(\mathcal{S}) \to \mathrm{Mark}(Y).$$

For each marking $\mu$ on $\mathcal{S}$, choose a pants decomposition $b$ of $Y$ such that $b$ has minimal intersection with $\pi_Y(\mu)$. We extend $b$ to a marking $\nu = \Pi_Y(\mu)$ on $Y$ as follows. For each curve $\alpha \in b$, choose transversal $t_\alpha$ in $Y$ such that $d_\alpha(t_\alpha, \mu)$ is minimal. The marking $\nu = \{(\alpha, t_\alpha) : \alpha \in b\}$ will be called an *induced marking* of $\mu$ on $Y$, and it is well-defined up to a bounded number of choices. It follows from Lemmas 2.4.2 and 2.5.4 that for any marking $\mu$, any non-annular domain $Y \subset \mathcal{S}$, and any domain $Z \subset Y$,

$$d_Z(\mu, \Pi_Y(\mu)) \leq M, \tag{2}$$

where $M$ depends only on $\mathcal{S}$.

DEFINITION 2.5.6 (Relative marking extension). *Let $\mu \in \mathrm{Mark}(\mathcal{S})$ and $c$ be a multicurve on $\mathcal{S}$. We extend $c$ to a marking $\mu' \in \mathrm{Mark}(\mathcal{S})$ relative to $\mu$ as follows. For each non-annular domain $Y$ in $\mathcal{S} \setminus c$, choose an induced marking $\Pi_Y(\mu)$ on $Y$. Then for each curve $\alpha \in c$, choose a transversal $t_\alpha$ with minimal $d_\alpha(t_\alpha, \mu)$. The union of $\{(\alpha, t_\alpha) : \alpha \in c\}$ with the set of induced markings $\Pi_Y(\mu)$ forms a marking $\mu' \in \mathrm{Mark}(\mathcal{S})$ which is well-defined up to a bounded number of choices.*

The following is an immediate consequence of our construction.

LEMMA 2.5.7. *Let $c$ be a multicurve on $\mathcal{S}$, $\mu$ any marking, and $\mu'$ an extension of $c$ relative to $\mu$. For any proper domain $Z \subset \mathcal{S}$, if $Z$ is contained in an essential component of $S \setminus c$, or if $Z$ is a curve in $c$, then*

$$d_Z(\mu', \mu) \leq M,$$

*where $M$ depends only on $\mathcal{S}$.*

*Proof.* For any curve $\alpha$ in $c$, the transversal $t_\alpha$ to $\alpha$ in $\mu'$ was chosen to be uniformly close to $\pi_\alpha(\mu)$. Thus $d_\alpha(\mu', \mu)$ is uniformly bounded by a constant depending on $\mathcal{S}$. Now suppose $Z \subseteq Y$ where $Y$ is a component of $\mathcal{S} \setminus c$. By construction $\pi_Y(\mu') = \mathrm{base}(\Pi_Y(\mu))$. Thus, by (2), $d_Z(\mu', \mu) = d_Z(\Pi_Y(\mu), \mu)$ is also uniformly bounded by a constant depending only on $\mathcal{S}$. □

**2.6 Hierarchies.** In the previous section, we introduced the marking graph $\mathrm{Mark}(\mathcal{S})$ which is quasi-isometric to $\mathcal{MCG}(\mathcal{S})$. In this section, we will introduce the theory of hierarchies, which is useful for constructing efficient paths in $\mathrm{Mark}(\mathcal{S})$. These paths are naturally associated to efficient representations of elements in $\mathcal{MCG}(\mathcal{S})$ in terms of the generators, thus justifying $\mathrm{Mark}(\mathcal{S})$ as a good combinatorial model for $\mathcal{MCG}(\mathcal{S})$.

The idea of hierarchies is to associate to every pair of markings a family of geodesics in curve complexes that behave well with subsurface projections. In order for the theory to work, we need to impose a condition on geodesics in curve complexes called *tightness*. Let $Y \subseteq \mathcal{S}$ be a domain. A *tight geodesic* $g$ in $\mathcal{C}(Y)$ is a sequence $\{v_0, \ldots, v_n\}$ of simplices in $\mathcal{C}(Y)$, such that any sequence of vertices in $g$ is a geodesic in $\mathcal{C}(Y)$ in the usual sense, and $v_{i-1} \cup v_{i+1}$ fill a subsurface $Z \subset Y$ such that $\partial Z = v_i$. We remark that the original definition [MM00, Definition 4.2] consists of more information.

It is a theorem of Masur–Minsky that any two points in $\mathcal{C}(Y)$ is connected by at least one and at most finitely many tight geodesics [MM00, Lemma 4.5 and Corollary 6.14]. Henceforth, a geodesic in a curve complex will always mean a tight geodesic. By an abuse of notation, we will refer to $v_i$'s as vertices of $g$. We will say the *length* of $g$ is $n$, and write $|g| = n$. We will say $Y$ is the *domain* or *support* of $g$, and write $D(g) = Y$. We will sometimes use the notation $[v_0, v_n]$ to mean any geodesic from $v_0$ to $v_n$ in $\mathcal{C}(Y)$. Since $\mathcal{C}(Y)$ is $\delta$-hyperbolic, all (finitely many) geodesics from $v_0$ to $v_n$ are fellow-travelers.

We now briefly sketch the definition of a hierarchy. For a complete definition (see [MM00, Definition 4.4]). A *hierarchy* on $\mathcal{S}$ is a collection $H$ of geodesics such that each geodesic $g \in H$ is supported on some domain $Y \subseteq \mathcal{S}$, with a distinguished *main geodesic* $g_H = [v_0, v_n]$ supported on $\mathcal{S}$, together with some additional structure and satisfying certain conditions which we now highlight. A hierarchy $H$ comes equipped with a pair of markings $I(H)$ and $T(H)$ on $\mathcal{S}$, called the *initial marking* and the *terminal marking* of $H$, respectively, such that $v_0 \subseteq \text{base}(I(H))$ and $v_n \subseteq \text{base}(T(H))$. We will usually assume $I(H)$ and $T(H)$ are complete clean marking on $\mathcal{S}$. One of the key technical conditions of a hierarchy is called *subordinacy*. Roughly, given a geodesic $g$ in $\mathcal{C}(\mathcal{S})$, one can inductively construct a hierarchy $H$ with $g = g_H$. For each vertex $v_i$ in $g_H$, the vertices $v_{i-1}$ to $v_{i+1}$ are contained in some component $Z$ of $\mathcal{S} \setminus v_i$. The geodesic $h = [v_{i-1}, v_{i+1}]$ in $\mathcal{C}(Z)$ will be an element of $H$ and is subordinate to $g_H$. One can continue this process with all vertices of $g_H$ and then with $h$ and so on.

We list some properties of hierarchies below, after the following definition.

DEFINITION 2.6.1 (Component domain). *Given a non-annular domain $Y \subseteq \mathcal{S}$, and a multicurve $c$ on $Y$, we say $Z$ is a component domain of $(Y, c)$ if $Z$ is either an essential component of $Y \setminus c$ or $Z$ is a curve in $c$.*

**Theorem 2.6.2.** *The following statements hold for hierarchies.*

1. (*Existence*) *Given any markings $\mu$ and $\nu$ on $\mathcal{S}$, there exists a hierarchy $H$ with $I(H) = \mu$ and $T(H) = \nu$* [MM00, Theorem 4.6].
2. (*Uniqueness of geodesics*) *For any hierarchy $H$, if $h, h' \in H$ have $D(h) = D(h')$, then $h = h'$* [MM00, Theorem 4.7].
3. (*Completeness*) *For every geodesic $h \in H$ and vertex $v \in h$, if $Y$ is a component domain of $(D(h), v)$, then $Y$ is domain for a geodesic $k \in H$* [MM00, Theorem 4.20].

We will sometimes denote an hierarchy from $\mu$ to $\nu$ by $H(\mu, \nu)$. The following lemma explains the relationship between a geodesic $h \in H$ and the projection of $I(H)$ and $T(H)$ to $D(h)$.

LEMMA 2.6.3 [MM00, Lemma 6.2]. *There exist constants $M_1 > M_2$, depending only on $\mathcal{S}$, such that if $H$ is any hierarchy in $\mathcal{S}$ and*

$$d_Y(I(H), T(H)) \geq M_2$$

*for a subsurface $Y$ in $\mathcal{S}$, then $Y$ is a domain for a geodesic $h \in H$.*

*Conversely, if $h \in H$ is any geodesic with $Y = D(h)$, then $h$ fellow travels any geodesic from $\pi_Y(I(H))$ to $\pi_Y(T(H))$ in $\mathcal{C}(Y)$ with a uniform constant. In particular,*

$$||h| - d_Y(I(H), T(H))| \leq M_1.$$

For any pair of markings $\mu, \nu \in \mathrm{Mark}(\mathcal{S})$, we will call a domain $Y$ a *large link* for $\mu$ and $\nu$ if $d_Y(\mu, \nu) \geq M_2$.

The theorem below summarizes two results that are vital to this paper. To simplify the statements we introduce some notations that we will adopt for the rest of the paper. Below, $\mathfrak{a}$ and $\mathfrak{b}$ represent quantities such as distances or lengths, and $k$ and $c$ are constants that depend only on $\mathcal{S}$ (unless otherwise noted).

**Notations 2.6.4.**   1. If $\mathfrak{a} \leq k\mathfrak{b} + c$, we say $a$ is *coarsely bounded* by $b$, and write $\mathfrak{a} \prec \mathfrak{b}$.

2. If $\dfrac{1}{k}\mathfrak{b} - c \leq \mathfrak{a} \leq k\mathfrak{b} + c$, we say $a$ is *coarsely equal* to $b$, and write $\mathfrak{a} \asymp \mathfrak{b}$.

By the *length* $|H|$ of a hierarchy $H$ we will mean $|H| = \sum_{h \in H} |h|$. In the following, the coarse equality on the left is [MM00, Theorem 6.10]. The coarse equality on the right is called the distance formula [MM00, Theorem 6.12].

**Theorem 2.6.5.**   *There exists a constant $L_0$ depending only on $S$ such that, for any $L \geq L_0$ and any $\mu, \nu \in \mathrm{Mark}(\mathcal{S})$ and any hierarchy $H = H(\mu, \nu)$,*

$$|H| \asymp d_{\mathrm{Mark}(\mathcal{S})}(\mu, \nu) \asymp \sum_{\substack{Y \subseteq \mathcal{S} \\ d_Y(\mu,\nu) \geq L}} d_Y(\mu, \nu).$$

*On the right, the constants involved in $\asymp$ depend on $L$.*

Fix a generating set $\Lambda$ for $\mathcal{MCG}(\mathcal{S})$. To realize the quasi-isometry between $\mathrm{Mark}(\mathcal{S})$ and $\mathcal{MCG}(\mathcal{S})$, we fix a base marking $\mu_B$ in $\mathrm{Mark}(\mathcal{S})$. Then $d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f\mu_B) \asymp |f|$, with constants depending only on $\mu_B$ and $\Lambda$. The following is an immediate consequence of Theorem 2.6.5.

COROLLARY 2.6.6 [MM00, Theorem 7.1]. *Let $\mu_B \in \mathrm{Mark}(\mathcal{S})$ be a fixed base marking. For any element $f \in \mathcal{MCG}(\mathcal{S})$,*

$$|H(\mu_B, f\mu_B)| \asymp |f|.$$

Let $Y \subset \mathcal{S}$ be a proper non-annular domain. There is a coarse embedding $\mathrm{Mark}(Y) \xrightarrow{j} \mathrm{Mark}(\mathcal{S})$ obtained as follows. Fix a marking in each essential component of $\mathcal{S} \setminus Y$ and a transversal to each curve in $\partial Y$. The map $j$ sends the marking $\nu \in \mathrm{Mark}(Y)$ in an obvious way so that $\partial Y \subseteq \mathrm{base}(j(\nu))$ and for all $\nu_1, \nu_2 \in \mathrm{Mark}(Y)$,

$$d_{\mathrm{Mark}(Y)}(\nu_1, \nu_2) \asymp d_{\mathrm{Mark}(\mathcal{S})}(j(\nu_1), j(\nu_2)). \tag{3}$$

Equation (3) follows from the distance formula and one can make the coarse constants independent of $Y$ and $j$.

**2.7  Slices.**    The connection between paths in $\mathrm{Mark}(\mathcal{S})$ and hierarchies come from *slices* of a hierarchy. The following definition comes from [MM00, §5].

DEFINITION 2.7.1 (Slices). *A (complete) slice of a hierarchy $H$ is a set $\tau$ of pointed geodesics $(h, v)$ in $H$, i.e. $h \in H$ and $v$ is a vertex of $h$, satisfying the following properties:*

*(S1) Any geodesic $h$ of $H$ appears at most once in $\tau$.*
*(S2) There is a distinguished pair, the bottom pair, $(g_H, b)$ of $\tau$.*
*(S3) For every $(k, w) \in \tau$ other than the bottom pair, $D(k)$ is a component domain of $(D(h), v)$ for some $(h, v) \in \tau$.*
*(S4) Given $(h, v) \in \tau$, for every component domain $Y$ of $(D(h), v)$ there is a pair $(k, w) \in \tau$ with $D(k) = Y$.*

The *initial slice* $\tau_0$ of $H$ is one where every pair $(h, v) \in \tau$ has $v$ the first vertex of $h$. In particular, the main geodesic $g_H$ and its initial vertex is a pair in $\tau_0$, and $\tau_0$ can be constructed inductively using the axioms of slices. Similarly, the *terminal slice* of $H$ is defined.

To any slice $\tau$ we can associate a complete marking $\mu_\tau$ as follows. First, let $\mu$ be the marking with

$$\mathrm{base}(\mu) = \{v : (h, v) \in \tau \text{ and } D(h) \text{ is not an annulus}\}.$$

For each base curve $\alpha$, if $(k, t) \in \tau$ is such that $k$ is a geodesic in $\mathcal{C}(\alpha)$, then let $t$ be the transversal to $\alpha$ in $\mu$. The marking $\mu$ is complete but not necessarily clean. Any clean marking $\mu_\tau$ compatible with $\mu$ will be called a *compatible* marking with $\tau$. By Lemma 2.5.1, the number of choices for $\mu_\tau$ is bounded. Note that $I(H)$ and $T(H)$ are, respectively, compatible markings with the initial and terminal slice of $H$. We will call any marking compatible with some slice in a hierarchy $H$ a *hierarchal marking* of $H$.

Given any slice $\tau$ in $H$, there is a notion of (*forward*) *elementary move* on $\tau$ which is roughly moving a vertex $v$ of some pair $(h, v) \in \tau$ forward by one step in the geodesic $h$ to obtain a new slice $\tau'$. We write $\tau \to \tau'$. (See [MM00, §5] for a precise definition.) If $\mu$ and $\mu'$ are compatible marking with $\tau$ and $\tau'$, then by Masur and Minsky [MM00, Lemma 5.5], $d_{\mathrm{Mark}(\mathcal{S})}(\mu, \mu') \prec 1$. We will write $\mu \to \mu'$ to mean any path in $\mathrm{Mark}(\mathcal{S})$ connecting $\mu$ to $\mu'$. To prove $|H| \asymp d_{\mathrm{Mark}(\mathcal{S})}(I(H), T(H))$, Masur and Minsky [MM00] established the existence of a *resolution* of $H$, which is a sequence of forward elementary moves

$$\tau_0 \to \cdots \to \tau_n,$$

where $\tau_0$ is the initial slice and $\tau_n$ is the terminal slice of $H$. For each $\tau_i$ in the resolution, let $\mu_i$ be a compatible marking with $\tau_i$. The corresponding path in $\mathrm{Mark}(\mathcal{S})$

$$I(H) = \mu_0 \to \cdots \to \mu_n = T(H)$$

is a quasi-geodesic with uniform constants, and $d_{\mathrm{Mark}(\mathcal{S})}(I(H), T(H)) \asymp n \asymp |H|$. A fact in [Min10] that we will sometime need is that, for any slice $\tau$ in $H$, there is a resolution of $H$ containing $\tau$.

The following statements are true for hierarchal markings and follow from [MM00].

LEMMA 2.7.2. *Let $H$ be a hierarchy. If $\mu \in \mathrm{Mark}(\mathcal{S})$ is a hierarchal marking of $H$, then*

$$d_{\mathrm{Mark}(\mathcal{S})}(I(H), \mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, T(H)) \prec d_{\mathrm{Mark}(\mathcal{S})}(I(H), T(H)). \tag{4}$$

*There exists a constant $M$ depending only on $\mathcal{S}$ such that for any domain $Y \subseteq \mathcal{S}$,*

$$d_Y(I(H), \mu) + d_Y(\mu, T(H)) \leq d_Y(I(H), T(H)) + M. \tag{5}$$

*Proof.* Let $g$ be a quasi-geodesic in $\mathrm{Mark}(\mathcal{S})$ containing $\mu$ coming from a resolution of $H$. By Masur and Minsky [MM00], $g$ is a quasi-geodesic from $I(H)$ to $T(H)$ with uniform constants, hence (4) holds.

Given $Y \subseteq \mathcal{S}$, let $\pi_Y(g)$ be the projection of $g$ to $\mathcal{C}(Y)$ (project each vertex of $g$ to $\mathcal{C}(Y)$). The projection $\pi_Y(g)$ is a quasi-geodesic in $\mathcal{C}(Y)$ with uniform constant. By hyperbolicity of $\mathcal{C}(Y)$, $\pi_Y(g)$ stays uniformly close to any geodesic connecting $\pi_Y(I(H))$ and $\pi_Y(T(H))$ of $\pi_Y(g)$. Thus there exists a constant $M_Y$ such that

$$d_Y(I(H), \mu) + d_Y(\mu, T(H)) \leq d_Y(I(H), T(H)) + M_Y.$$

Since there are only finitely many subsurfaces of $\mathcal{S}$ up to homeomorphism, the constant $M = \max_Y \{M_Y\}$ depends only on $\mathcal{S}$ and achieves (5). $\qquad\square$

**2.8  Time order.**     The geodesics or domains of geodesics in a hierarchy $H$ satisfy a partial order $<_t$, called *time order*. We refer to [MM00, §4.6] for the definition. The idea comes from the observation that the vertices of a geodesic $g$ are linearly ordered: $v_i < v_j$ if $i < j$. Combining this observation with the subordinacy structure on $H$, one can try to order a pair of geodesics $h, h' \in H$. In the following, we summarize some main results and state some useful consequences of time order.

**Theorem 2.8.1** [MM00, Lemma 4.18 and 4.19]*.   There exists a relation $<_t$, called time-order, on domains of geodesics in $H$ such that:*

- *The relation $<_t$ is a strict partial order.*
- *If $h$ and $h'$ are geodesics in $H$ such that $Y = D(h)$ and $Z = D(h')$ interlock, then either $Y <_t Z$ or $Z <_t Y$.*
- *If $Y \subset Z$, then $Y$ and $Z$ are not time-ordered.*
- *If $Y$ and $Z$ lie in different component domains of $(D(m), v)$, for some geodesic $m$ in $H$ and $v \in m$, then $Y$ and $Z$ are not time-ordered.*

Note that the ambiguous case is when $D(h)$ and $D(h')$ are disjoint; sometimes they are time-ordered and sometimes not. The issue of disjoint domains will come up in this paper.

The constant $M_1$ of Lemma 2.6.3 can be chosen so that following hold.

LEMMA 2.8.2 [MM00, Lemma 6.11]. *Let $H$ be a hierarchy. Suppose $Y$ and $Z$ are domains for geodesics in $H$ such that $Y$ and $Z$ interlock. If $Y <_t Z$, then $d_Y(\partial Z, T(H)) \leq M_1$ and $d_Z(I(H), \partial Y) \leq M_1$.*

Using slices, the constant $M_1$ can be chosen so the following version of Lemma 2.8.2 also holds.

LEMMA 2.8.3. *With the same hypothesis as above. There exists a hierarchy marking $\nu$ such that $d_Y(\nu, T(H)) \leq M_1$ and $d_Z(I(H), \nu) \leq M_1$.*

*Proof.* By assumption, both $Y$ and $Z$ are domains for geodesics for a hierarchy $H$ with $Y <_t Z$. Let $k \in H$ be the geodesic supported on $Y$ and let $w \in k$ be the terminal vertex of $k$. Let $\tau$ be a slice $\tau$ with $(k, w) \in \tau$ (such $\tau$ exists by Minsky [Min10, Lemma 5.8]). Let $\nu$ a hierarchal marking compatible with $\tau$. Since $k$ is supported on $Y$, by definition of a slice, there exists some pair $(h, u) \in \tau$ such that $Y$ is a component domain of $(D(h), v)$. By definition of a compatible marking, we have $\partial Y \subseteq \text{base}(\nu)$, which implies that, by Lemma 2.8.2, $d_Z(I(H), \nu)$ is uniformly bounded. Since $w$ is the terminal vertex of $k$, any resolution of $H$ containing $\tau$ does not pass through $Y$ from $\tau$ to $T(H)$. Thus, $d_Y(\nu, T(H))$ is also uniformly bounded. This finishes the proof of the lemma. □

We may choose $M_1$ so the following also holds:

LEMMA 2.8.4 [BM08, Lemma 1]. *With the same hypothesis as above. For any marking $\mu \in \text{Mark}(\mathcal{S})$, either $d_Y(\mu, T(H)) \leq 2M_1$ or $d_Z(I(H), \mu) \leq 2M_1$.*

**2.9 Separating Marking.** The following definition and lemma do not explicitly appear in [MM00]. Although the lemma is a direct consequence of hierarchies, we offer a brief sketch of its proof.

DEFINITION 2.9.1 (Separating marking). *Let $H$ be a hierarchy. A slice $\tau$ is called a separating slice if for every pair $(h, v) \in \tau$, with $h \neq g_H$, has the property that $v$ is the terminal vertex of $h$. We remark that once the bottom pair $(g_H, b)$ is fixed, then the separating slice $\tau$ containing $(g_H, b)$ is uniquely determined by the axioms of slices. In particular, if $b$ is the terminal vertex of $g_H$, then $\tau$ is the terminal slice of $H$. If $\tau$ is a separating slice containing $(g_H, b)$, then any marking $\mu$ compatible with $\tau$ is called a separating marking at $b$.*

The constant $M_1$ of Lemma 2.6.3 can be chosen so that the following hold.

LEMMA 2.9.2. *Let $H$ be a hierarchy. Let $b$ be any vertex in $g_H$ and let $\mu$ be a separating marking at $b$. Then for any proper domain $Y \subset \mathcal{S}$, either $d_Y(I(H), \mu) \leq M_1$ or $d_Y(\mu, T(H)) \leq M_1$.*

*Proof.* We may assume $Y \subset \mathcal{S}$ has $d_Y(I(H), T(H)) > M_1$. Since $M_1 \geq M_2, Y$ is a domain for a geodesic $h_Y \in H$. Without a loss of generality, we may assume $Y$ is a component domain of $(g_H, c)$, for some $c$ in $g_H$. If $c$ appears before $b$ along $g_H$, then

$d_Y(\mu, T(H)) \leq M_1$. Similarly, if $c$ appears after $b$ along $g_H$, then $d_Y(I(H), \mu) \leq M_1$. Both of these facts can be seen as a consequence of Lemma 2.8.3. The remaining case is $b = c$. In this case, the separating slice containing $(g_H, b)$ must contain $(h_Y, v)$, where $v$ is the terminal vertex of $h_Y$. Therefore, it must be that $d_Y(\mu, T(H)) \leq M_1$. □

REMARK 2.9.3. In our definition of separating slice, the preference for terminal vertices is arbitrary. Lemma 2.9.2 would remain true if we allowed only initial vertices or a mixture of initial and terminal.

**2.10  Collecting constants.**    For the rest of the paper, we will fix the following set of constants.

Let $M_0$ be the constant of Theorem 2.4.3. Let $L_0$ be the constant of Theorem 2.6.5. Let $M_1$ and $M_2$ be the constants coming from Lemma 2.6.3. We will also fix one constant $M_3$ for Lemmas 2.4.2, 2.5.1, Eq. (2) and Lemmas 2.5.7, 2.7.2. We may assume $M_1 \geq M_2, M_3$. In addition, since up to homeomorphism there are only finitely many subsurfaces of $\mathcal{S}$, we can choose a hyperbolicity constant $\delta$ that works for all $\mathcal{C}(Z), Z \subseteq \mathcal{S}$.

# 3  Two Technical Lemmas

This section contains some technical results about finite order mapping classes.

To prove L.B.C. property, we need to understand the geometry of the action of finite order mapping classes on $\mathrm{Mark}(\mathcal{S})$. The first observation is that finite order elements act on $\mathrm{Mark}(\mathcal{S})$ with coarse fixed points. We will eventually prove that the action has the property that the translation distance of a finite order element $f$, or $d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f\mu_B)$ where $\mu_B$ is the base marking, is coarsely bounded by the distance from $\mu_B$ to the fixed point sets of $f$. In other words, finite order elements of $\mathcal{MCG}(\mathcal{S})$ act *elliptically* on $\mathrm{Mark}(\mathcal{S})$.

In this section, we consider what happens if a fixed point $\mu$ of $f$ is far from $\mu_B$ relative to the translation distance of $f$. By the distance formula, there must be some $X \subseteq S$ such that $d_X(\mu_B, \mu)$ is large relative to $d_X(\mu_B, f\mu_B)$. With some additional conditions, $X$ will be called a *bad domain* for $\mu$ and we will prove a structure theorem for the set of bad domains in a hierarchy $H(\mu_B, \mu)$. In the next section, we will use this structure theorem to construct a coarse fixed point of $f$ close to $\mu_B$ relative to the translation distance of $f$. From there, we can derive L.B.C. property for finite order mapping classes by a standard argument.

**3.1  Fixed points and symmetric points.**    We state some useful facts about finite order mapping classes below.

LEMMA 3.1.1. *There are finitely many conjugacy classes of finite order elements in $\mathcal{MCG}(\mathcal{S})$.*

COROLLARY 3.1.2. *There exists a constant $N$, depending only on $\mathcal{S}$, such any finite order element $f \in \mathcal{MCG}(\mathcal{S})$ has $\mathrm{order}(f) \leq N$.*

DEFINITION 3.1.3. *Let $f \in \mathcal{MCG}(\mathcal{S})$ be of finite order. We define the set of $r$-fixed points of $f$ as*

$$\mathrm{Fix}_r(f) = \{\mu \in \mathrm{Mark}(\mathcal{S}) : d_{\mathrm{Mark}(\mathcal{S})}(\mu, f\mu) \leq r\}.$$

*Also, define the set of $r$-symmetric points for $f$ to be*

$$\widetilde{\mathrm{Fix}}_r(f) = \{\mu \in \mathrm{Mark}(\mathcal{S}) : d_Y(\mu, f\mu) \leq r, \forall Y \subseteq \mathcal{S}\}.$$

LEMMA 3.1.4. *There exists a constant $R_1$ depending only on $\mathcal{S}$ such that $\mathrm{Fix}_{R_1}(f) \neq \emptyset$ and $\widetilde{\mathrm{Fix}}_{R_1}(f) \neq \emptyset$, for any finite order element $f \in \mathcal{MCG}(\mathcal{S})$.*

*Proof.* Choose any $\mu \in \mathrm{Mark}(\mathcal{S})$ and let $R_f = d_{\mathrm{Mark}(\mathcal{S})}(\mu, f\mu)$. If $g = \omega f \omega^{-1}$ for some $\omega \in \mathcal{MCG}(\mathcal{S})$, then $d_{\mathrm{Mark}(\mathcal{S})}(\omega\mu, g\omega\mu) = d_{\mathrm{Mark}(\mathcal{S})}(\mu, f\mu) \leq R_f$. Thus $\mathrm{Fix}_{R_f}(g) \neq \emptyset$ for all $g$ in the conjugacy class of $f$. Using Lemma 3.1.1, we can let $R_1$ be the maximum of the constants $R_f$ ranging over all conjugacy classes. This gives $\mathrm{Fix}_{R_1}(f) \neq \emptyset$ for all finite order element $f \in \mathcal{MCG}(\mathcal{S})$. Using the distance formula, we can choose $R_1$ so that $\widetilde{\mathrm{Fix}}_{R_1}(f) \neq \emptyset$ as well. □

Henceforth, we will fix $R_1$ to be the minimal constant satisfying Lemma 3.1.4.

REMARK 3.1.5. We can describe the geometry of the subset $\mathrm{Fix}_{R_1}(f) \subset \mathrm{Mark}(\mathcal{S})$. By Nielsen Realization, any finite order element $f \in \mathcal{MCG}(\mathcal{S})$ can be realized as an isometry of a hyperbolic surface $X$ [Ker83]. The quotient $\bar{X} = X/f$ is an orbifold. One can coarsely identify $\mathrm{Fix}_R(f)$ with $\mathrm{Mark}(\bar{X})$. The map $X \to \bar{X}$ is a (branched) covering map. By Rafi and Schleimer [RS09], the lifting of $\mathrm{Mark}(\bar{X})$ to $\mathrm{Mark}(X) = \mathrm{Mark}(\mathcal{S})$ is a quasi-isometric embedding.

**3.2  An example.**    Before proceeding to the first technical lemma, let's discuss a motivating example. The following refers to Figure 1.
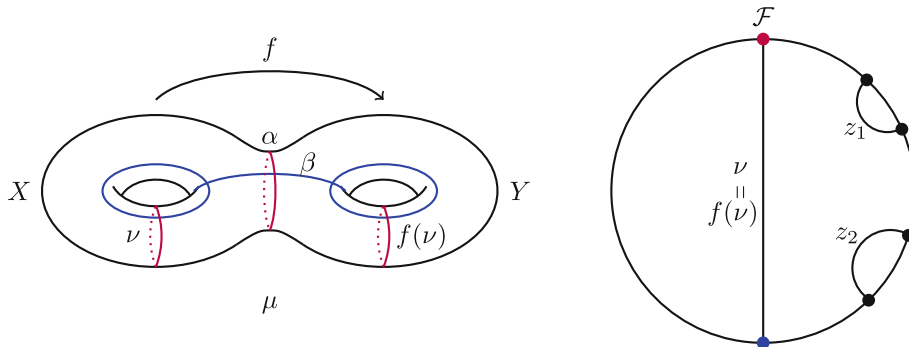


Figure 1: On the *left*, the curves represent a 0-fixed point $\mu$ for the order two element $f \in \mathcal{MCG}(\mathcal{S}_2)$ that permutes the holes of $\mathcal{S}_2$. On the *right* is the Farey graph $\mathcal{F}$ which is isomorphic to $\mathcal{C}(X)$ and $\mathcal{C}(Y)$; $\mathcal{C}(X)$ and $\mathcal{C}(Y)$ are identified via $f$. Markings on $X$ or $Y$ correspond to edges in $\mathcal{F}$. Here, $\Pi_X(\mu) = \nu$ and $\Pi_Y(\mu) = f(\nu)$ represent the same edge in $\mathcal{F}$. The base marking $\mu_B$ (which is not drawn on the *left*) has $\Pi_X(\mu_B) = z_1$ and $\Pi_Y(\mu_B) = z_2$

Consider the closed surface $\mathcal{S}_2$ of genus two. Let $f$ be the mapping class of order two that permutes the two holes of $S_2$. Let $\alpha$ be the separating curve in $\mathcal{S}$ indicated in Figure 1. Let $X$ and $Y$ be the pair of once-punctured tori in $S_2$ with boundary $\partial X = \alpha = \partial Y$ The map $f$ permutes $X$ and $Y$. Using this and the fact that $X$ and $Y$ are disjoint, we can construct a family of coarse fixed points of $f$ as follows (see Remark 3.1.5). Since $X$ is a once-punctured torus, $\mathcal{C}(X)$ is homeomorphic to the Farey graph $\mathcal{F}$. Via the map $f$, we can also identify $\mathcal{C}(Y)$ with $\mathcal{F}$. After this identification, markings on $X$ or $Y$ correspond to edges in $\mathcal{F}$. Choose any marking $\nu$ on $X$. By the action of $f$, we get a marking $f(\nu)$ on $Y$, which is represented by the same edge in $\mathcal{F}$. Choose a transverse curve $\beta$ to $\alpha$ so that $\mu = \nu \cup f(\nu) \cup \alpha \cup \beta$ is a clean marking. Since $f(\mu) = \mu$, $\mu$ is a fixed point of $f$.

See Fig. 1 for a concrete example of a fixed point $\mu$ of $f$ from this construction, where we have color-coded the curves so that the red or vertical curves represent the base curves of a marking $\mu$ and the blue or horizontal curves represent the transversal curves of $\mu$. In the example, let $\nu_1$ be the marking on $X$ obtained by the $(0, 1)$ and $(1, 0)$ curves. The marking $\mu = \nu \cup f(\nu) \cup \alpha \cup \beta$ is a 0-fixed point of $f$.

Consider a base marking $\mu_B$ constructed as follows. For simplicity, we will assume $\alpha$ is a base curve of $\mu_B$. Choose two edges $z_1$ and $z_2$ in $\mathcal{F}$ that are very far apart. We will let $z_1$ be the marking in $X$ and $z_2$ be the marking in $Y$. Now choose a transverse curve $\beta'$ to $\alpha$ so that $\mu_B = z_1 \cup z_2 \cup \alpha \cup \beta'$ is clean. Since $z_1$ and $z_2$ are far, $\mu_B$ is itself not a coarse fixed point of $f$. Let $\mathcal{Z}_B$ be the collection of domains on which $d_Z(\mu_B, f\mu_B) \geq L_0$, where $L_0$ is the constant of Theorem 2.6.5. Since $\alpha$ is a base curve of $\mu_B$, if $Z \in \mathcal{Z}_B$ then either $Z \subseteq X$ or $Z \subseteq Y$. Also, since $d_Z(\mu_B, f\mu_B) = d_{f(Z)}(\mu_B, f\mu_B)$, if $Z \in \mathcal{Z}_B$ then $f(Z) \in \mathcal{Z}_B$. Finally, since $d_X(\mu_B, f\mu_B) = d_{\mathcal{F}}(z_1, z_2)$ is large, $X$ (and $Y$) is in $\mathcal{Z}_B$. By Theorem 2.6.5,

$$d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f\mu_B) \asymp \sum_{Z \in \mathcal{Z}_B} d_Z(\mu_B, f\mu_B).$$

To find a fixed point $\mu$ of $f$ "close" to $\mu_B$,

$$d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \mu) \prec d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f\mu_B), \tag{6}$$

consider the following construction. Let $g$ be a geodesic in $\mathcal{F}$ connecting base$(z_1)$ and base$(z_2)$ (note that the convex hull of base$(z_1)$ and base$(z_2)$ is a finite set of geodesics). Let $\nu$ be any edge in $g$, which we will regard as a marking in $X$. Let $\beta$ be a transverse curve to $\alpha$ so that $d_\alpha(\beta, \beta')$ is uniformly bounded and $\mu = \nu \cup f(\nu) \cup \alpha \cup \beta$ is clean (Lemma 2.5.1). We show $\mu$ is "close" to $\mu_B$. By assumption, $d_\alpha(\mu_B, \mu)$ is uniformly bounded. Since $\alpha$ is contained in both $\mu_B$ and $\mu$, if $Z$ is any domain that contains $\alpha$ or is crossed by $\alpha$, we have $d_Z(\mu_B, \mu) \leq 4$. For any $Z \subseteq X$, we have (ignoring some addictive errors)

$$d_Z(\mu_B, \mu) = d_Z(z_1, \nu) \leq d_Z(z_1, f(z_2)) = d_Z(\mu_B, f\mu_B).$$

Similarly, for any $Z \subseteq Y$,

$$d_Z(\mu_B, \mu) = d_Z(z_2, f(\nu)) \leq d_Z(z_2, f(z_1)) = d_Z(\mu_B, f\mu_B).$$

Let $\mathcal{Z}_\mu$ be the set of domains on which $d_Z(\mu_B, \mu) \geq L_0$. By the above computations, if $Z \in \mathcal{Z}_\mu$, then $Z \in \mathcal{Z}_B$. Thus we have

$$
\begin{aligned}
d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \mu_1) &\asymp \sum_{Z \in \mathcal{Z}_\mu} d_Z(\mu_B, \mu_1) \\
&\leq \sum_{Z \in \mathcal{Z}_\mu} d_Z(\mu_B, f\mu_B) \\
&\leq \sum_{Z \in \mathcal{Z}_B} d_Z(\mu_B, f\mu_B) \\
&\asymp d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f\mu_B).
\end{aligned}
$$

Hence $\mu$ satisfies (6). We emphasize that, by varying the choice of the edge in $g$, we obtain a family of fixed points of $f$ "close" to $\mu_B$.

In the following and in the subsequent section, we generalize this example. The general situations could be much more complicated; for instance, the assumption that $\mu_B$ contains $\alpha$ simplified the example quite a bit. The reason why our construction worked is because, in every domain $Z \subseteq \mathcal{S}, d_Z(\mu_B, \mu) \leq d_Z(\mu_B, f\mu_B)$. Thus, if a coarse fixed point $\mu$ is not "close" to $\mu_B$, then there should be some $Z \subseteq \mathcal{S}$ on which $d_Z(\mu_B, \mu) > d_Z(\mu_B, f\mu_B)$. This is the motivation behind Definition 3.3.1 of a bad domain $Z$ for $\mu$ (the actual definition contains a slightly stronger condition). In our construction of coarse fixed points, we relied heavily on the fact that $X$ and $Y$ are disjoint and $X = f(Y)$. In general, we will also try to look for a domain $X$ such that $\{f^i(X)\}$ are all pairwise disjoint. The structure result for bad domains, Lemma 3.3.4, shows that if $X$ is a bad domain for $\mu$, then $X$ and its orbits are all disjoint and are all (essentially) bad domains for $\mu$. In Section 4, we will show that, when a coarse fixed point $\mu$ of $f$ does not have any bad domains, then $\mu$ will be the desired marking "close" to $\mu_B$ (Proposition 4.1.1). Otherwise, we will show how to use the disjointness result of a bad domain $X$ and its orbits to construct a new coarse fixed point of $f$ "closer" to $\mu_B$ (see Section 4.2). Finitely many iterations of this construction will lead to a desired coarse fixed point of $f$ "close" to $\mu_B$ (see Section 4.3).

**3.3 Bad domains and the first technical lemma.** We remark that Lemma 3.1.1 does not a priori help us with L.B.C. property as each conjugacy class has infinitely many elements, but it will play an essential role later.

In the following, let $\mu_B \in \mathrm{Mark}(\mathcal{S})$ be the fixed base marking. We recall notations of Section 2.10 and let $R_1$ be the minimal constant satisfying Lemma 3.1.4. Set

$$\Theta = 6M_1 + 4\delta. \tag{7}$$

Let $f \in \mathcal{MCG}(\mathcal{S})$ be of finite order. Fix $N$ to be the smallest constant satisfying Corollary 3.1.2. For any proper domain $X \subset \mathcal{S}$, let $L_X = L_X(f)$ be the integer such that $f^{L_X+1}$ is the first return map of $f$ to $X$. Note that for any $X, L_X < \mathrm{order}(f) \leq N$.

DEFINITION 3.3.1 (Bad domains). Let $R \geq R_1$ and let $\mu \in \mathrm{Mark}(\mathcal{S})$ be any marking. We say a domain $X \subseteq \mathcal{S}$ is a *R-bad domain* for $\mu$ (and $f$) if

- For $X = \mathcal{S}$,

$$d_{\mathcal{S}}(\mu_B, \mu) > d_{\mathcal{S}}(\mu_B, f\mu_B) + R. \tag{8}$$

- For $X \neq \mathcal{S}$,

$$d_X(\mu_B, \mu) > 2N \left( \max_{0 \leq i \leq L_X} \{ d_{f^i(X)}(\mu_B, f\mu_B) \} \right) + NR + \Theta. \tag{9}$$

Denote by $\Omega(\mu, R, f)$ or $\Omega(\mu, R)$ the set of all $R$-bad domains for $\mu$ (and $f$). Note that if $R' > R$, then $\Omega(\mu, R') \subseteq \Omega(\mu, R)$. We remark that the constant "$2N$" in the definition of bad domains will not play a role until the next section.

DEFINITION 3.3.2 (Partial order on $\Omega(\mu, R)$). *We endow $\Omega(\mu, R)$ with a partial order "$\lhd$" following these rules. Let $X, Y \in \Omega(\mu, R)$, and let $H = H(\mu_B, \mu)$ be a fixed hierarchy.*

*(O1) If $\xi(X) < \xi(Y)$, then $X \lhd Y$. In particular, if $\mathcal{S} \in \Omega(\mu, R)$, then $\mathcal{S}$ is the maximal element.*
*(O2) If $\xi(X) = \xi(Y)$ and $X <_t Y$ in $H(\mu_B, \mu)$, then $Y \lhd X$.*

DEFINITION 3.3.3 (Complexity of $\Omega(\mu, R)$). *If $\Omega(\mu, R)$ is non-empty, then the complexity of the maximal element in $\Omega(\mu, R)$ is called the complexity of $\Omega(\mu, R)$, denoted by $\xi(\mu, R)$. The minimal complexity over all subsurfaces of $\mathcal{S}$ is $-1$, coming from an annulus. For convenience, we will let $\xi(\emptyset) = -2$. We make the trivial observations that*

$$\Omega(\mu', R') \subseteq \Omega(\mu, R) \implies \xi(\mu', R') \leq \xi(\mu, R).$$

The following is a consequence of the definition of $R$-bad domains for $\mu$ and $f$, if $\mu$ happens to be a $R$-symmetric point for $f$.

LEMMA 3.3.4 (Structure of bad domains). *Let $R \geq R_1$ and let $\mu \in \widetilde{\mathrm{Fix}}_R(f)$. If $X \in \Omega(\mu, R)$ and $X \neq \mathcal{S}$, then*

$$X, f(X), \ldots f^{L_X}(X)$$

*are all domains for geodesics in $H(\mu_B, \mu)$ and are all pairwise disjoint.*

*Proof.* Note that $f^{-n}(X) = f^{L_X + 1 - n}(X)$. We will prove, by induction on $n$, that

$$X, f^{-1}(X), \ldots, f^{-n}(X) = f^{L_X + 1 - n}(X)$$

satisfy the conclusion of the lemma for $n = 0, \ldots, L_X$. Our assumption is that $X \in \Omega(\mu, R)$ and $X \neq \mathcal{S}$. In particular, this means $d_X(\mu_B, \mu) > \Theta > M_2$, so $X$ is a domain for a geodesic in $H(\mu_B, \mu)$. This concludes the base case. Let's now assume

$$X, f^{-1}(X), \ldots, f^{-n+1}(X)$$

are all domains for geodesics in $H(\mu_B, \mu)$ and are all pairwise disjoint. We will show $f^{-n}(X)$ supports a geodesic in $H(\mu_B, \mu)$. Recursively, we have

$$
\begin{aligned}
d_{f^{-n}(X)}&(\mu_B, \mu) \\
&= d_{f^{-n+1}(X)}(f\mu_B, f\mu) \\
&\geq d_{f^{-n+1}(X)}(\mu_B, \mu) - d_{f^{-n+1}(X)}(\mu_B, f\mu_B) - d_{f^{-n+1}(X)}(\mu, f\mu) \\
&> d_{f^{-n+2}(X)}(f\mu_B, f\mu) - d_{f^{-n+1}(X)}(\mu_B, f\mu_B) - R \\
&\qquad\vdots \\
&\geq d_X(\mu_B, \mu) - \left( \sum_{i=1}^{n} d_{f^{-n+i}(X)}(\mu_B, f\mu_B) \right) - nR \\
&\geq d_X(\mu_B, \mu) - n \left( \max_{1 \leq i \leq n} \{ d_{f^{-n+i}(X)}(\mu_B, f\mu_B) \} \right) - nR \\
\text{By } (9) \ &> (N-n) \left( \max_{1 \leq i \leq L_X+1} \{ d_{f^{-n+i}(X)}(\mu_B, f\mu_B) \} \right) + (N-n)R + \Theta.
\end{aligned}
$$

Since $N > L_X \geq n$, we have in particular

$$
d_{f^{-n}(X)}(\mu_B, \mu) > \Theta > M_2.
$$

Therefore, $f^{-n}(X)$ supports a geodesic in $H(\mu_B, \mu)$.

Now let's prove $f^{-n}(X)$ is disjoint from each $X, \ldots, f^{-n+1}(X)$. Observe that $f^{-n}(X)$ and $f^{-i}(X)$ are disjoint if and only if $f^{-n+i}(X)$ and $X$ are disjoint. Hence it is enough to show $f^{-n}(X)$ and $X$ are disjoint. By way of contradiction, let's assume $X$ and $f^{-n}(X)$ are not disjoint. The two domains have the same complexity so they must interlock. They both support geodesics in $H(\mu_B, \mu)$ so, by Theorem 2.8.1, they are time-ordered. We have two cases.

The first case is $X <_t f^{-n}(X)$. As in Lemma 2.8.3, we may choose a hierarchal marking $\nu$ for $H(\mu_B, \mu)$ such that

$$
d_X(\nu, \mu) \leq M_1 \quad \text{and} \quad d_{f^{-n}(X)}(\mu_B, \nu) \leq M_1, \tag{10}
$$

By the triangle inequality,

$$
\begin{aligned}
d_X(\mu_B, f^n \nu) &\leq d_X(\mu_B, f^n \mu_B) + d_X(f^n \mu_B, f^n \nu) \\
&\leq \left( \sum_{j=0}^{n-1} d_X(f^j \mu_B, f^{j+1} \mu_B) \right) + d_{f^{-n}(X)}(\mu_B, \nu) \\
&= \left( \sum_{j=0}^{n-1} d_{f^{-j}(X)}(\mu_B, f\mu_B) \right) + d_{f^{-n}(X)}(\mu_B, \nu) \\
&\leq n \left( \max_{0 \leq j \leq n-1} \{ d_{f^{-j}(X)}(\mu_B, f\mu_B) \} \right) + M_1. \tag{11}
\end{aligned}
$$

Using the triangle inequality again, along with (9) and (11), we have

$$d_X(f^n\nu, \mu) \geq d_X(\mu_B, \mu) - d_X(\mu_B, f^n\nu) > 2M_1.$$

Therefore, by Lemma 2.8.4,

$$d_{f^{-n}(X)}(\mu_B, f^n\nu) \leq 2M_1.$$

Now we consider $d_X(\mu_B, f^{in}\mu)$. By iterating the argument we obtain inductively, for every $i \geq 0$,

$$\begin{aligned}
d_X(\mu_B, f^{(i+1)n}\nu) &\leq d_X(\mu_B, f^n\mu_B) + d_X(f^n\mu_B, f^{(i+1)n}\nu) \\
&\leq n\left(\max_{0\leq j\leq n-1}\{d_{f^{-j}(X)}(\mu_B, f\mu_B)\}\right) + d_{f^{-n}(X)}(\mu_B, f^{in}\nu) \\
&\leq n\left(\max_{0\leq j\leq n-1}\{d_{f^{-j}(X)}(\mu_B, f\mu_B)\}\right) + 2M_1.
\end{aligned} \tag{12}$$

Since there is some $i$ for which $f^{in}$ is the identity map, inequality (12) must also hold for $d_X(\mu_B, \nu)$. With this fact coupled with the first half of (10), we derive the following violation to $X \in \Omega(\mu, R)$:

$$\begin{aligned}
d_X(\mu_B, \mu) &\leq d_X(\mu_B, \nu) + d_X(\nu, \mu) \\
&\leq n\left(\max_{0\leq j\leq L_X}\{d_{f^{-j}(X)}(\mu_B, f\mu_B)\}\right) + 2M_1 + d_X(\nu, \mu) \\
&\leq N\left(\max_{0\leq j\leq L_X}\{d_{f^{-j}(X)}(\mu_B, f\mu_B)\}\right) + 3M_1.
\end{aligned}$$

Thus, it is not possible for $X <_t f^{-n}(X)$.

To eliminate the second case $f^{-n}(X) <_t X$, we argue similarly. Now choose a marking $\nu$ such that

$$d_X(\mu_B, \nu) \leq M_1 \quad \text{and} \quad d_{f^{-n}(X)}(\nu, \mu) \leq M_1. \tag{13}$$

Then, using the fact that $\mu \in \widetilde{\text{Fix}}_R(f)$ (in the last step below), we have

$$\begin{aligned}
d_X(f^n\nu, \mu) &\leq d_X(f^n\nu, f^n\mu) + d_X(\mu, f^n\mu) \\
&\leq d_{f^{-n}(X)}(\nu, \mu) + \sum_{i=0}^{n-1} d_{f^{-i}(X)}(\mu, f\mu) \\
&\leq M_1 + nR.
\end{aligned}$$

Thus,

$$d_X(\mu_B, f^n\nu) \geq d_X(\mu_B, \mu) - d_X(f^n\nu, \mu) > 2M_1.$$

By Lemma 2.8.4,

$$d_{f^{-n}(X)}(f^n\nu, \mu) \leq 2M_1.$$

As above, we iterate the argument on powers of $f^n$. For all $i \geq 0$, we have

$$
\begin{aligned}
d_X(f^{(i+1)n}\nu, \mu) &\leq d_X(f^{(i+1)n}\nu, f^n\mu) + d_X(\mu, f^n\mu) \\
&\leq d_{f^{-n}(X)}(f^{in}\nu, \mu) + nR \\
&\leq 2M_1 + nR.
\end{aligned}
$$

This eventually leads to the contradiction

$$
\begin{aligned}
d_X(\mu_B, \mu) &\leq d_X(\mu_B, \nu) + d_X(\nu, \mu) \\
&\leq M_1 + 2M_1 + nR \\
&\leq 3M_1 + nR.
\end{aligned}
$$

We conclude $X$ and $f^{-n}(X)$ must be disjoint. $\qquad\square$

**3.4  Second technical lemma.**    In the following, we prove another technical result, which has a similar conclusion as Lemma 3.3.4, but it is based on different assumptions. Its purpose is for the situation when the main surface $\mathcal{S}$ is a bad domain for an $R$-fixed point $\mu$ of a finite order mapping class $f$ (see Proposition 4.2.1). In this situation, we need to cook up a set of base curves for a new fixed point of $f$ that is closer to the base marking $\mu_B$ in $\mathcal{C}(\mathcal{S})$. Lemma 3.4.1 starts this process by finding a subsurface whose orbit under $f$ are all pairwise disjoint. Furthermore, the boundary curves of these subsurfaces form a multicurve that is closer to $\mu_B$ in $\mathcal{C}(\mathcal{S})$ than the base curves of $\mu$.

The proof of Lemma 3.4.1 will be similar to that of Lemma 3.3.4. We will provide the necessary details.

In the following, let $R$ be any constant and let $\mu \in \widetilde{\mathrm{Fix}}_R(f)$ for a finite order mapping class $f$. Let $H(\mu_B, f\mu_B)$ and $H(\mu_B, \mu)$ be hierarchies. Let $[v_B, f(v_B)]$ and $[v_B, v]$ be the corresponding main geodesics in $\mathcal{C}(\mathcal{S})$. For any domain $Y \subset \mathcal{S}$, we adopt the notation $[u, v]_Y$ to mean the line segment $[\pi_Y(u), \pi_Y(v)]$ in $\mathcal{C}(Y)$. For any two sets $A, B \subset \mathcal{C}(\mathcal{S})$, let $\mathrm{Dist}_{\mathcal{S}}(A, B) = \min_{\substack{u \in A \\ v \in B}} d_{\mathcal{S}}(u, v)$.

LEMMA 3.4.1. *There exists a constant $\Delta$, depending only on $R$, such that the following holds. Suppose $b$ is a vertex on $[v_B, v]$ with the property that*

$$
\mathrm{Dist}_{\mathcal{S}}([b, f(b)], [v_B, f(v_B)]) \geq 4.
$$

*Let $\mu'$ be a separating marking at $b$. Then whenever a subsurface $Y \subset \mathcal{S}$ has the property that*

$$
\mathrm{Dist}_{\mathcal{S}}(\partial Y, [b, f(b)]) \leq 1 \quad \text{and} \quad d_Y(\mu', f\mu') > \Delta,
$$

*then*

$$
Y, f(Y), \dots, f^{L_Y}(Y)
$$

*are all domains for geodesics in $H(\mu_B, \mu)$ and are all mutually disjoint.*

*Proof.* We claim the constant

$$\Delta = (2N + 1)M_0 + (2N + 1)R + 10M_1$$

works. Let $Y \subset \mathcal{S}$ satisfy the criteria of the lemma. As in Lemma 3.3.4, we will prove by induction on $n$ that

$$Y, f^{-1}(Y), \ldots, f^{-n}(Y) = f^{L_Y + 1 - n}(Y)$$

satisfy the lemma for $n = 0 \ldots, L_Y$.

Let's first show $Y$ supports a geodesic in $H(\mu_B, \mu)$. We will be considering $H(f\mu_B, f\mu)$ in parallel. Note that $f\mu'$ will be a separating marking at $f(b)$ in $H(f\mu_B, f\mu)$. Consider the following pair of quadrilaterals in $\mathcal{C}(Y)$:

$$Q_1 = [\mu_B, f\mu_B]_Y \cup [f\mu_B, f\mu']_Y \cup [\mu_B, \mu']_Y \cup [\mu', f\mu']_Y,$$

and

$$Q_2 = [\mu', f\mu']_Y \cup [\mu', \mu]_Y \cup [\mu, f\mu]_Y \cup [f\mu', f\mu]_Y.$$

By our assumption, $d_Y(\mu', f\mu') > \Delta$. By the triangle inequality, at least one of three other segments of $Q_1$ is long:

$$[\mu_B, \mu']_Y, \quad [\mu_B, f\mu_B]_Y, \quad [f\mu_B, f\mu']_Y.$$

Similarly, at least one of the following segments of $Q_2$ is long:

$$[\mu', \mu]_Y, \quad [\mu, f\mu]_Y, \quad [f\mu', f\mu]_Y.$$

Since $d_Z(\mu, f\mu) \leq R$ for all $Z \subseteq \mathcal{S}$, in $Q_2$ the situation reduces to either

$$d_Y(\mu', \mu) > \frac{\Delta - R}{2} > NM_0 + NR + 5M_1, \tag{14}$$

or

$$d_Y(f\mu', f\mu) > NM_0 + NR + 5M_1. \tag{15}$$

If (14) holds, then by the fact that $\mu'$ is a separating marking at $b$ (Lemma 2.9.2),

$$d_Y(\mu_B, \mu') \leq M_1.$$

Applying the triangle inequality yields

$$\begin{aligned} d_Y(\mu_B, \mu) &\geq d_Y(\mu', \mu) - d_Y(\mu_B, \mu') \\ &> (NM_0 + NR + 5M_1) - M_1. \end{aligned}$$

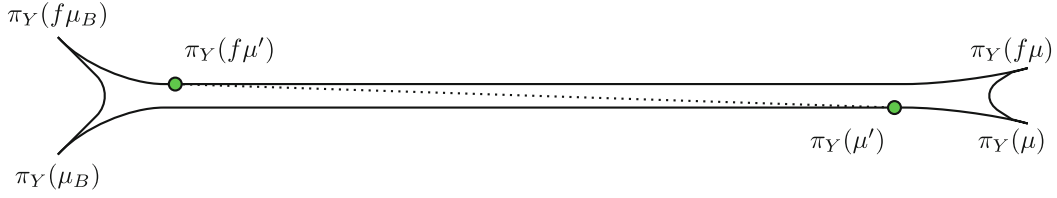Therefore, (14) implies $Y$ supports a geodesic in $H$. So we may assume (15) holds.

Figure 2: The quadrilaterals $Q_1$ and $Q_2$ in $\mathcal{C}(Y)$

In $Q_1$, the assumption on $\partial Y$ forces $\text{Dist}_{\mathcal{S}}(\partial Y, [\mu_B, f\mu_B]) > 1$. In other words, every vertex in $[\mu_B, f\mu_B]$ crosses $Y$. Theorem 2.4.3 applies and $d_Y(\mu_B, f\mu_B) \leq M_0$. The situation is reduced to either

$$d_Y(\mu_B, \mu') > \frac{\Delta - M_0}{2} > NM_0 + NR + 5M_1 \tag{16}$$

or

$$d_Y(f\mu_B, f\mu') > NM_0 + NR + 5M_1. \tag{17}$$

It is not possible for (15) and (17) to occur simultaneously, as that would mean both

$$d_Y(f\mu_B, f\mu') > M_1 \quad \text{and} \quad d_Y(f\mu', f\mu) > M_1,$$

violating $f\mu'$ a separating marking. So (16) must hold. As above, we must then have

$$d_Y(\mu_B, \mu) \geq d_Y(\mu_B, \mu') - d_Y(\mu', \mu)$$
$$> (NM_0 + NR + 5M_1) - M_1.$$

Therefore, in all cases, $Y$ must support a geodesic in $H(\mu_B, \mu)$. See Figure 2 for a schematic picture of $Q_1$ and $Q_2$. Note that the conclusion of the base case always resulted in

$$d_Y(\mu_B, \mu) > NM_0 + NR + 4M_1. \tag{18}$$

By induction,

$$Y, f^{-1}(Y), \ldots, f^{-n+1}(Y)$$

are all domains for geodesics in $H(\mu_B, \mu)$ and are all mutually disjoint. Let's now prove $f^{-n}(Y)$ supports a geodesic in $H(\mu_B, \mu)$. Since

$$\text{Dist}_{\mathcal{S}}([v_B, f(v_B)], [b, f(b)]) \geq 4 \quad \text{and} \quad \text{Dist}_{\mathcal{S}}(\partial Y, [b, f(b)]) \leq 1,$$

the disjointness condition will imply

$$\text{Dist}_{\mathcal{S}}([v_B, f(v_B)], \partial f^{-i}(Y)) \geq 2,$$

for all $i = 0, \ldots, n - 1$. By Theorem 2.4.3,

$$d_{f^{-i}(Y)}(\mu_B, f\mu_B) \leq M_0. \tag{19}$$

Coupling this fact with (18) (in the last step below), we have

$$
\begin{aligned}
d_{f^{-n}(Y)}&(\mu_B, \mu) \\
&= d_{f^{-n+1}(Y)}(f\mu_B, f\mu) \\
&\geq d_{f^{-n+1}(Y)}(\mu_B, \mu) - d_{f^{-n+1}(Y)}(\mu_B, f\mu_B) - d_{f^{-n+1}(Y)}(\mu, f\mu) \\
&\geq d_{f^{-n+1}(Y)}(\mu_B, \mu) - M_0 - R \\
&\quad\ \vdots \\
&\geq d_Y(\mu_B, \mu) - nM_0 - nR \\
&> (N-n)M_0 + (N-n)R + 4M_1.
\end{aligned}
$$

Since $N > L_Y \geq n$, the above in particular implies

$$
d_{f^{-n}(Y)}(\mu_B, \mu) > M_1 \geq M_2.
$$

So $f^{-n}(Y)$ supports a geodesic in $H(\mu_B, \mu)$.

We now want to show $Y, \ldots, f^{-n+1}(Y), f^{-n}(Y)$ are all pairwise disjoint. Using the action of $f$ and the assumption that $Y, \ldots, f^{-n+1}(Y)$ are pairwise disjoint, we see that $f^{-n}(Y)$ is disjoint with each $Y, \ldots, f^{-n+1}(Y)$ if and only if $f^{-n}(Y)$ and $Y$ are disjoint. If $Y$ and $f^{-n}(Y)$ are not disjoint, then they are time-ordered in $H(\mu_B, \mu)$. The two different cases of time-ordering of $Y$ and $f^{-n}(Y)$ will both lead to a contradiction. The argument is very similar to the one given in Lemma 3.4.1. We will quickly give the argument in the case that $Y <_t f^{-n}(Y)$ and omit the argument in the second case.

Suppose $Y <_t f^{-n}(Y)$. Let $\nu$ be a hierarchy marking $H(\mu_B, \mu)$ such that

$$
d_Y(\nu, \mu) \leq M_1 \quad \text{and} \quad d_{f^{-n}(Y)}(\mu_B, \nu) \leq M_1, \tag{20}
$$

as in Lemma 2.8.4. Then

$$
\begin{aligned}
d_Y(\mu_B, f^n\nu) &\leq d_Y(\mu_B, f^n\mu_B) + d_Y(f^n\mu_B, f^n\nu) \\
&\leq \left( \sum_{j=0}^{n-1} d_Y(f^j\mu_B, f^{j+1}\mu_B) \right) + d_{f^{-n}(Y)}(\mu_B, \nu) \\
&\leq \left( \sum_{j=0}^{n-1} d_{f^{-j}(Y)}(\mu_B, f\mu_B) \right) + M_1 \\
&\leq nM_0 + M_1.
\end{aligned}
$$

Using (18) and the triangle inequality, we have

$$
d_Y(f^n\nu, \mu) \geq d_Y(\mu_B, \mu) - d_Y(\mu_B, f^n\nu) > 2M_1.
$$

Therefore, by Lemma 2.8.4,

$$
d_{f^{-n}(Y)}(\mu_B, f^n\nu) \leq 2M_1.
$$

By considering powers of $f^n\nu$ inductively, we have

$$d_Y(\mu_B, f^{(i+1)n}\nu) \leq d_Y(\mu_B, f^n\mu_B) + d_Y(f^n\mu_B, f^{(i+1)n}\nu)$$
$$\leq nM_0 + d_{f^{-n}(Y)}(\mu_B, f^{in}\nu)$$
$$\leq nM_0 + 2M_1.$$

This is true for every $i \geq 0$. Since $N > L_Y \geq n$, using (20), we have

$$d_Y(\mu_B, \mu) \leq d_Y(\mu_B, \nu) + d_Y(\nu, \mu) \leq nM_0 + 3M_1,$$

contradicting (18). The case of $f^{-n}(Y) <_t Y$ will lead to a similar contradiction. This concludes the proof of the lemma.                                    □

## 4  L.B.C. Property for Finite Order Mapping Classes

The heart of this section is to prove Theorem C of the introduction, restated below. Let $R_1$ be the fixed constant of Lemma 3.1.4 and let $\mu_B$ be the fixed base marking.

**Theorem 4.0.2.**  *There exists a constant $R \geq R_1$, depending only on $\mu_B$, such that any finite order $f \in \mathcal{MCG}(\mathcal{S})$ has a marking $\mu \in \widetilde{\mathrm{Fix}}_R(f)$ with*

$$d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \mu) \prec d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f\mu_B).$$

Assuming Theorem 4.0.2, we can derive L.B.C. property for finite order mapping classes by a standard argument, following [BH99]. We first state and prove the following corollary of Theorem 4.0.2, which reduces L.B.C. property for finite order mapping classes to a finite problem.

COROLLARY 4.0.3. *There exists a finite set $\Gamma \subset \mathcal{MCG}(\mathcal{S})$ such that, for every finite order $f \in \mathcal{MCG}(\mathcal{S})$, there exists $\omega \in \mathcal{MCG}(\mathcal{S})$ such that $\omega^{-1} f \omega \in \Gamma$ and $|\omega| \prec |f|$.*

*Proof.* By enlarging $R$ if necessary, we may rephrase Theorem 4.0.2 in terms of fixed points: there exists $R$ depending only on $\mu_B$ such that any finite order mapping class $f$ has a marking $\mu \in \mathrm{Fix}_R(f)$ with

$$d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \mu) \prec d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f\mu_B). \tag{21}$$

We construct the set $\Gamma$ as follows. Let $D$ be the diameter of $\mathrm{Mark}(\mathcal{S})$ modulo the action of $\mathcal{MCG}(\mathcal{S})$. ($D$ is finite since the action of $\mathcal{MCG}(\mathcal{S})$ on $\mathrm{Mark}(\mathcal{S})$ is cofinite). Set

$$\Gamma = \{g \in \mathcal{MCG}(\mathcal{S}) : d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, g\mu_B) \leq 2D + R, g \text{ finite order}\}.$$

The action of $\mathcal{MCG}(\mathcal{S})$ on $\mathrm{Mark}(\mathcal{S})$ is proper, thus $\Gamma$ is a finite set. We show $\Gamma$ satisfies the other properties as well.

Let $f \in \mathcal{MCG}(\mathcal{S})$ be of finite order. Let $\mu$ be a $R$-fixed point of $f$ closest to $\mu_B$. Since the action of $\mathcal{MCG}(\mathcal{S})$ on $\mathrm{Mark}(\mathcal{S})$ is cofinite, there exists $\omega \in \mathcal{MCG}(\mathcal{S})$ such that $d_{\mathrm{Mark}(\mathcal{S})}(\omega\mu_B, \mu) \leq D$. Then $\omega^{-1}f\omega \in \Gamma$, since

$$
\begin{aligned}
d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \omega^{-1}f\omega\mu_B) &= d_{\mathrm{Mark}(\mathcal{S})}(\omega\mu_B, f\omega\mu_B) \\
&\leq d_{\mathrm{Mark}(\mathcal{S})}(\omega\mu_B, \mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, f\mu) + d_{\mathrm{Mark}(\mathcal{S})}(f\mu, f\omega\mu_B) \\
&\leq d_{\mathrm{Mark}(\mathcal{S})}(\omega\mu_B, \mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, f\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, \omega\mu_B) \\
&\leq 2D + R.
\end{aligned}
$$

Moreover, by (21) we have

$$
\begin{aligned}
|\omega| &\prec d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \omega\mu_B) \\
&\leq d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \mu) + d_{\mathrm{Mark}(\mathcal{S})}(\omega\mu_B, \mu) \\
&\prec d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f\mu_B) + D \\
&\prec |f|. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square
\end{aligned}
$$

COROLLARY 4.0.4 (L.B.C. property for finite order mapping classes). *If $f, g \in \mathcal{MCG}(\mathcal{S})$ are conjugate finite order mapping classes, then there is a conjugating element $\omega \in \mathcal{MCG}(\mathcal{S})$ with*

$$
|\omega| \prec |f| + |g|.
$$

*Proof.* Let $\Gamma \subset \mathcal{MCG}(\mathcal{S})$ be the finite set of Theorem 4.0.2. The content of Theorem 4.0.2 is that $\Gamma$ contains at least one and at most finitely many representatives for each conjugacy class of a finite order mapping class. Furthermore, each finite order $f$ can be conjugated into $\Gamma$ by a conjugating element whose word length is proportional to $|f|$. The result follows after picking a conjugating element for each pair of elements in $\Gamma$ of the same conjugacy class. $\qquad\qquad \square$

The proof of Theorem 4.0.2 will occupy the rest of the section. The main observation is that if $\mu_1 \in \widetilde{\mathrm{Fix}}_{R_1}(f)$ does not have any $R_1$–bad domains, then $\mu_1$ satisfies the statement of Theorem 4.0.2 (see Proposition 4.1.1). If $\mu_1$ does have a $R_1$-bad domain $Y$, then we can construct a marking $\mu_2 \in \widetilde{\mathrm{Fix}}_{R_2}(f)$, where $R_2$ depends only on $R_1$, such that $Y \notin \Omega(\mu_2, R_2)$. We will call this the base step of the proof. Ideally, we would like $\Omega(\mu_2, R_2)$ to be strictly smaller than $\Omega(\mu_1, R_1)$, but the situation is a bit more complicated. In trying to improve $\mu_1$ in $Y$, we may create new bad domains, but we will have control over what they are in relation to $Y$. Although the set of bad domains is not necessarily decreasing, applying the base step in the right way will guarantee a decrease in the complexity of $\Omega(\mu_2, R_2)$ from that of $\Omega(\mu_1, R_1)$. By iterating this process, we will produce a sequence of symmetric points for $f$ such that the complexities of the sets of bad domains are monotonically decreasing. This process must stop to produce an $R$-symmetric point $\mu$ for $f$ with no bad domains. Since the maximal complexity of any set of bad domains is the complexity of $\mathcal{S}$, the

process of achieving $\mu$ terminates after a bounded number of steps. This serves to ensure the constant $R$ will depend only on $R_1$ (and $\mu_B$).

The rest of the section is organized as follows. We first prove in Proposition 4.1.1 that no bad domain indeed implies Theorem 4.0.2. Then the base step is dealt with in Section 4.2. There are two propositions, Propositions 4.2.1 and 4.2.2, associated to the base step, depending on whether the bad domain is the main surface or a proper subsurface. This is where our work in Section 3 will come in. In Section 4.3, we will explain how to use the base step to reduce the complexity of the set of bad domains. The precise statement is Proposition 4.3.1. The section will conclude with Corollary 4.3.2 which makes precise how the process terminates after a bounded number of steps.

## 4.1    No bad domains

PROPOSITION 4.1.1 (No bad domains). *Let $\mu \in \widetilde{\mathrm{Fix}}_R(f)$ where $R \geq R_1$. If $R$ is a constant depending only on $R_1$ such that $\Omega(\mu, R) = \emptyset$, then*

$$d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \mu) \prec d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f\mu_B).$$

*In other words, $\mu$ and $R$ satisfy Theorem 4.0.2.*

*Proof.* The assumption $\Omega(\mu, R) = \emptyset$ means that for every $X \subseteq \mathcal{S}$, there exists $i_X$ such that

$$d_X(\mu_B, \mu) \leq 2N d_{f^{i_X}(X)}(\mu_B, f\mu_B) + NR + \Theta.$$

Let $L_0$ be the constant of the distance formula, Theorem 2.6.5. Let

$$\Phi = \{X \subseteq \mathcal{S} : d_X(\mu_B, \mu) \geq 2NL_0 + NR + \Theta\},$$

and

$$\Psi = \{Y \subseteq \mathcal{S} : d_Y(\mu_B, f\mu_B) \geq L_0\}.$$

Then there is a map $\Phi \to \Psi$ sending $X \mapsto f^{i_X}(X)$. This map has multiplicity at most the order of $f$, which is bounded by $N$. Therefore,

$$\begin{aligned}
d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \mu) &\asymp \sum_{X \in \Phi} d_X(\mu_B, \mu) \\
&\leq \sum_{X \in \Phi} 2N d_{f^{i_X}(X)}(\mu_B, f\mu_B) + NR + \Omega \\
&\prec \sum_{X \in \Phi} d_{f^{i_X}(X)}(\mu_B, f\mu_B) \\
&\leq N \sum_{Y \in \Psi} d_Y(\mu_B, f\mu_B) \\
&\asymp d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f\mu_B). \qquad \square
\end{aligned}$$

**4.2   Base step.**    We are now ready to state and prove the base step of the proof for Theorem 4.0.2. There are two cases to consider, which are Propositions 4.2.1 and 4.2.2. The proof of Proposition 4.2.1 will be essential for Proposition 4.2.2.

Let $\mu_B$ be the base marking in $\mathrm{Mark}(\mathcal{S})$, and recall the definition of $\xi(\Omega(\mu, R))$ as in Definition 3.3.3.

PROPOSITION 4.2.1 (Base step 1). *Given $R_I \geq \max\{2\delta + 4, R_1\}$ there exists a constant $R_O$ depending only on $R_I$ with the following property. Given $\mu_I \in \widetilde{\mathrm{Fix}}_{R_I}(f)$, if $\mathcal{S} \in \Omega(\mu_I, R_I)$, then there exists $\mu_O \in \mathrm{Mark}(\mathcal{S})$ satisfying the following properties:*

(P1) $\mu_O \in \widetilde{\mathrm{Fix}}_{R_O}(f)$.
(P2) $\Omega(\mu_O, R_O) \subsetneqq \Omega(\mu_I, R_I)$. *In addition,* $\mathcal{S} \notin \Omega(\mu_O, R_O)$, *and thus* $\xi(\mu_O, R_O) \subsetneqq \xi(\mu_I, R_I)$.

*Proof.* We have four hierarchies:

$$H(\mu_B, f\mu_B), \quad H(\mu_I, f\mu_I), \quad H(\mu_B, \mu_I), \quad H(f\mu_B, f\mu_I).$$

Consider the four main geodesics corresponding to the four hierarchies, forming a quadrilateral $Q$ in $\mathcal{C}(\mathcal{S})$:

$$[v_B, f(v_B)], \quad [v_I, f(v_I)], \quad [v_B, v_I], \quad [f(v_B), f(v_I)],$$

where $v_B$ and $v_I$ are base curves in $\mu_B$ and $\mu_I$, respectively. Our assumption is that $\mathcal{S} \in \Omega(\mu_I, R_I)$, so

$$d_{\mathcal{S}}(v_B, v_I) > d_{\mathcal{S}}(v_B, f(v_B)) + R_I.$$

Since $f$ acts on $\mathcal{C}(\mathcal{S})$ as an isometry, $d_{\mathcal{S}}(v_B, v_I) = d_{\mathcal{S}}(f(v_B), f(v_I))$, so $Q$ is $2\delta$-thin: every edge of $Q$ is contained in a $2\delta$-neighborhood of the other edges. The geodesics $[v_I, v_B]$ and $[f(v_I), f(v_B)]$ $2\delta$-fellow travel for awhile until $[v_I, v_B]$ begins fellow traveling $[v_B, f(v_B)]$. Choose the vertex $b$ on $[v_I, v_B]$ at the junction where this change takes place. After possibly moving $b$ toward $v_I$, by at most $2\delta + 4$ positions, we may assume the following properties for $b$ (see Figure 3):



Figure 3: The quadrilateral $Q$ in $\mathcal{C}(\mathcal{S})$

- $d_{\mathcal{S}}(b, f(b)) \leq 2\delta$, and
- $4 \leq \mathrm{Dist}_{\mathcal{S}}([b, f(b)], [v_B, f(v_B)]) \leq 6\delta + 4$.

By the triangle inequality, we have

$$d_{\mathcal{S}}(b, v_B) \leq d_{\mathcal{S}}(v_B, f(v_B)) + 8\delta + 4. \tag{22}$$

Now choose the separating marking $\mu$ in $H(\mu_B, \mu_I)$ at $b$. The proof divides into two cases. To describe these cases, consider any domain $Y \subset \mathcal{S}$ on which

$$\mathrm{Dist}_{\mathcal{S}}(\partial Y, [b, f(b)]) \leq 1, \tag{23}$$

and let

$$\Delta = (2N+1)M_0 + (2N+1)R_I + 10M_1.$$

**Case I.** Suppose $d_Y(\mu, f\mu) \leq \Delta$ for all $Y \subset \mathcal{S}$ satisfying (23). In this case, set $\mu_O = \mu$ and $R_O = \Delta$. We show $\mu_O$ and $R_O$ satisfy Proposition 4.2.1. First note that $\mathcal{S} \notin \Omega(\mu_O, R_O)$ by (22). Since $\mu_O$ is a hierarchal marking in $H(\mu_B, \mu_I)$, we also have, for all $Z \subset \mathcal{S}$,

$$d_Z(\mu_B, \mu_O) \leq d_Z(\mu_B, \mu_I) + M_3.$$

Since $M_3 < M_1 < R_O$, this verifies (P2).

To see property (P1), we consider $d_Z(\mu_O, f\mu_O)$ for three possibilities of $Z$.

(a1) If $d_{\mathcal{S}}(\partial Z, [b, f(b)]) \leq 1$, then

$$d_Z(\mu_O, f\mu_O) \leq \Delta = R_O.$$

(a2) If $d_{\mathcal{S}}(\partial Z, [b, f(b)]) > 1$, then every vertex of $[b, f(b)]$ cuts $Z$. By Theorem 2.4.3,

$$d_Z(b, f(b)) \leq M_0 \implies d_Z(\mu_O, f\mu_O) \leq M_0 + 4.$$

(a3) If $Z = \mathcal{S}$, then by construction, $d_{\mathcal{S}}(b, f(b)) \leq 2\delta$, thus

$$d_{\mathcal{S}}(\mu_O, f\mu_O) \leq 2\delta + 4.$$

This ends the proof of the proposition in Case I.

**Case II.** Suppose there exists $Y$ with $d_Y(\mu, f\mu) > \Delta$ for some $Y \subset \mathcal{S}$ satisfying (23). In this case, Lemma 3.4.1 implies that $Y$ and all its orbits under $f$ are pairwise disjoint. Consider the multicurve

$$c = \partial Y \cup \partial f(Y) \cup \cdots \cup \partial f^{L_Y}(Y).$$

Let $\mu_O$ be a marking extension of $c$ relative to $\mu_I$, as in Definition 2.5.6. In particular, $c \subseteq \mathrm{base}(\mu_O)$. Set

$$R_O = \max\{R_I + 2M_3, 10\delta + 13\}.$$

Consider the following properties for $\mu_O$ and $R_O$.

(b1)   If $Z = \mathcal{S}$, then, since both $\mu_O$ and $f\mu_O$ contain $c$ as base curves,

$$d_{\mathcal{S}}(\mu_O, f\mu_O) \leq d_{\mathcal{S}}(\mu_O, c) + d_{\mathcal{S}}(c, f\mu_O) \leq 4. \tag{24}$$

Also, since

$$\mathrm{Dist}_{\mathcal{S}}(\partial Y, [b, f(b)]) \leq 1,$$

we have

$$
\begin{aligned}
d_{\mathcal{S}}(\mu_B, \mu_O) &\leq d_{\mathcal{S}}(v_B, \partial Y) + 4 \\
&\leq d_{\mathcal{S}}(v_B, b) + d_{\mathcal{S}}(b, f(b)) + 5 \\
&\leq (d_{\mathcal{S}}(v_B, f(v_B)) + 8\delta + 4) + 2\delta + 5 \\
&\leq d_{\mathcal{S}}(v_B, f(v_B)) + 10\delta + 9 \\
&\leq d_{\mathcal{S}}(\mu_B, f(\mu_B)) + 10\delta + 13.
\end{aligned} \tag{25}
$$

(b2)   If $Z \neq \mathcal{S}$, but some curve $\alpha$ in $c$ crosses $Z$, then

$$d_Z(\mu_O, f\mu_O) \leq d_Z(\mu_O, \alpha) + d_Z(\alpha, f\mu_O) \leq 8. \tag{26}$$

Furthermore, according to Lemma 3.4.1, $f^i(Y)$ all support a geodesic in $H(\mu_B, \mu_I)$, so there exists a slice of $H(\mu_B, \mu_I)$ containing $\alpha$. Therefore,

$$
\begin{aligned}
d_Z(\mu_B, \mu_O) &\leq d_Z(\mu_B, \alpha) + 4 \\
&\leq d_Z(\mu_B, \mu_I) + M_3 + 4.
\end{aligned} \tag{27}
$$

(b3)   If $Z \neq \mathcal{S}$ is such that $Z$ is disjoint from $c$ or $Z$ is curve in $c$, then we are in the situation of Lemma 2.5.7. Note that $f\mu_O$ is a marking extension of $c$ relative to $f\mu_I$.

$$
\begin{aligned}
d_Z(\mu_O, f\mu_O) &\leq d_Z(\mu_O, \mu_I) + d_Z(\mu_I, f\mu_I) + d_Z(f\mu_I, f\mu_O) \\
&\leq d_Z(\mu_I, f\mu_I) + 2M_3 \\
&\leq R_I + 2M_3 \\
&\leq R_O,
\end{aligned} \tag{28}
$$

and

$$d_Z(\mu_B, \mu_O) \leq d_Z(\mu_B, \mu_I) + d_Z(\mu_I, \mu_O) \leq d_Z(\mu_B, \mu_I) + M_3. \tag{29}$$

Using the analyses of (b1)–(b3), we verify properties (P1) and (P2) for $\mu_O$ and $R_O$. From (24), (26), (28), we have that, for any $Z \subseteq \mathcal{S}$, $d_Z(\mu_O, f\mu_O) \leq R_O$. Thus $\mu_O \in \widetilde{\mathrm{Fix}}_{R_O}(f)$ and (P1) is verified. To see (P2), first note that, by (25), we have

$\mathcal{S} \notin \Omega(\mu_O, R_O)$. Now, if $Z \in \Omega(\mu_O, R_O)$ where $Z \subset \mathcal{S}$, then $Z$ must be of case (b2) or (b3). In either case, using (27) or (29), we obtain

$$
\begin{aligned}
d_Z(\mu_B, \mu_I) \ &\geq d_Z(\mu_B, \mu_O) - (M_3 + 4) \\
&> 2N\left(\max_i\{d_{f^i(Z)}(\mu_B, f\mu_B)\}\right) + NR_O + \Theta - (M_3 + 4) \\
&> 2N\left(\max_i\{d_{f^i(Z)}(\mu_B, f\mu_B)\}\right) + N(R_I + 2M_3) + \Theta - (M_3 + 4) \\
&> 2N\left(\max_i\{d_{f^i(Z)}(\mu_B, f\mu_B)\}\right) + NR_I + \Theta.
\end{aligned}
$$

Therefore, $\Omega(\mu_O, R_O) \subset \Omega(\mu_I, R_I)$, establishing (P2). This finishes the proof of the proposition in Case II. □

Before we state the next proposition we will need some notations and definitions. Given proper domains $X, Y \subset \mathcal{S}$, let

$$\mathcal{U} = X \cup \cdots \cup f^{L_X}(X) \quad \text{and} \quad \mathcal{V} = Y \cup \cdots \cup f^{L_Y}(Y).$$

We will say $Y$ is *supported* on $\mathcal{S} \setminus \mathcal{U}$ if $Y$ lies in some component of $\mathcal{S} \setminus \mathcal{U}$. In the case that $X$ is not a curve, $Y$ can be a boundary curve of $f^i(X)$, for some $0 \leq i \leq L_X$. Note the symmetry in the definition: if $Y$ is supported on $\mathcal{S} \setminus \mathcal{U}$ then $X$ is supported on $\mathcal{S} \setminus \mathcal{V}$. Furthermore, if $Y$ is supported on $\mathcal{S} \setminus \mathcal{U}$, then so is $f^j(Y)$ for all $j = 0, \ldots L_Y$. Thus it makes sense to say that $\mathcal{U}$ and $\mathcal{V}$ are disjoint. Similarly, given $X_1 \ldots X_n \subset \mathcal{S}$ and let

$$\mathcal{U}_i = X_i \cup \cdots \cup f^{L_{X_i}}(X_i),$$

we will say $\mathcal{U}_1, \ldots \mathcal{U}_n$ are pairwise disjoint if, for all $1 \leq i, j \leq n$, $i \neq j$, $X_i$ is supported on $\mathcal{S} \setminus \mathcal{U}_j$.

PROPOSITION 4.2.2 (Base step 2). *Given $R_I \geq \max\{2\delta + 4, R_1\}$ there exists a constant $R_O$ depending only on $R_i$ with the following property. Given $\mu_I \in \widetilde{\mathrm{Fix}}_{R_I}(f)$ and suppose $\mathcal{S} \notin \Omega(\mu_I, R_I)$. If $\Omega(\mu_I, R_I)$ contains proper domains $X_1, \ldots X_n \subset \mathcal{S}$ such that $\mathcal{U}_1, \ldots, \mathcal{U}_n$ are pairwise disjoint, where $\mathcal{U}_i = X_i \cup \cdots \cup f^{L_{X_i}}(X_i)$, then there exists $\mu_O \in \mathrm{Mark}(\mathcal{S})$ satisfying the following properties:*

(Q1)   *$\mu_O \in \widetilde{\mathrm{Fix}}_{R_O}(f)$.*
(Q2)   *For $j = 1, \ldots n$, let $c_j = \partial X_i \cup \partial f(X_i) \cup \cdots \partial f^{L_{X_i}}(X_i)$. Then $c = \bigcup_j c_j \subseteq \mathrm{base}(\mu_O)$.*
(Q3)   *For all $j = 1, \ldots n$ and all $i = 0, \ldots L_{X_j}$, $f^i(X_j) \notin \Omega(\mu_O, R_O)$.*
(Q4)   *Suppose $Z \in \Omega(\mu_I, R_I)$ has the property that $Z$ interlocks $f^i(X_j)$, for some $0 \leq j \leq n$ and some $0 \leq i \leq L_{X_j}$. If $X_j <_t f^{-i}(Z)$ in $H(\mu_B, \mu_I)$, then $Z \notin \Omega(\mu_O, R_O)$.*
(Q5)   *If $Z \in \Omega(\mu_O, R_O)$ but $Z \notin \Omega(\mu_I, R_I)$, then $Z$ must be a subsurface of $f^i(X_j)$, for some $0 \leq j \leq n$ and $0 \leq i \leq L_{X_j}$. In particular, $\xi(Z) < \xi(X_j)$.*

REMARK 4.2.3. We briefly explain the statements in (Q1)–(Q5).

First note that, by Lemma 3.3.4, for each $j = 1, \ldots n$,

$$X_j, f(X_j), \ldots, f^{L_{X_j}}(X_j)$$

are all pairwise disjoint. Since the $\mathcal{U}_i$'s are assumed to be pairwise disjoint, the set $c = \bigcup_j c_j$ is a multicurve on $\mathcal{S}$, so property (Q2) makes sense. (Recall that if $X_j$ is a curve, then $\partial X_j = X_j$.)

Secondly, the assumption in (Q4) also makes sense. If $Z \in \Omega(\mu_I, R_I)$ interlocks $f^i(X_j)$, then $X_j$ and $f^{-i}(Z)$ interlock by the action of $f$. Since they both support geodesics in $H(\mu_B, \mu_I)$ (Lemma 3.3.4), they must be time-ordered. The proposition analyzes the case when $X_j <_t f^{-i}(Z)$.

The point of property (Q3) is that, if $X_j$ is a bad domain for $\mu_I$, then we can improve $\mu_I$ in $X_j, f(X_j), \ldots f^{L_{X_j}}(X_j)$ simultaneously. This process also eliminates all bad domains of type specified by (Q4). However, during this process, a new bad domain $Z$ that was not a bad domain for $\mu_I$ may have been created. Property (Q5) puts restrictions on such $Z$: namely, $Z$ must be a subsurface of some $f^i(X_j)$, so it has strictly smaller complexity than that of $X_j$. If $X_1, X_2, \ldots X_n$ are all curves, then in particular (Q5) implies such $Z$ cannot exist and thus $\Omega(\mu_O, R_O) \subset \Omega(\mu_I, R_I)$.

*Proof of Proposition* 4.2.2. We will first assume $n = 1$ and set $X = X_1$. We will construct a marking $\mu_O$ containing

$$c = \partial X \cup \cdots \cup \partial f^{L_X}(X).$$

as base curves, guaranteeing (Q2). The situation may seem similar to case II of Proposition 4.2.1, but to ensure (Q3), it will not be enough to construct $\mu_O$ by inducing $\mu_I$ on each $f^i(X)$. We will in fact need the full work of Proposition 4.2.1 to construct a marking on $X$. The action of $f$ will then extend this marking to each $f^i(X)$. We will consider two cases, when $X$ is a curve or when $X$ is a non-annular subsurface. The two cases are pretty much the same, but for clarity we treat them separately. After we explain how to construct $\mu_O$ and $R_O$ in each case, we will then check that they satisfy the proposition.

First suppose $X$ is a curve. On each component domain of $(S, c)$, we endow with the marking induced from $\mu_I$. To complete this into a marking, we need to pick a transversal to each $f^i(X)$. Much like in the proof of Proposition 4.2.1, we have a quadrilateral in $\mathcal{C}(X)$ formed by projecting the main geodesics $\mu_B, f\mu_B, \mu_I$, and $f\mu_I$ to $\mathcal{C}(X)$. Since the pair of geodesics $[\mu_B, \mu_I]_X$ and $[f\mu_B, f\mu_I]_X$ $2\delta$-fellow travel in $\mathcal{C}(X)$, we can find an element $b \in \mathcal{C}(X)$ such that

- $d_X(b, f(b)) \leq 2\delta$.
- $d_X(b, \mu_B) \leq d_X(\mu_B, f\mu_B) + 2\delta$.

Let $f^i(b)$ be the transversal to $f^i(X)$ and let $\mu_O$ be the associated clean marking on $S$. The correct constant will be $R_O = R_I + 2M_3$.

Now suppose $X$ is a non-annular domain. Let $F = f^{L_X+1} : X \to X$ be the first return map of $f$ to $X$. Set

$$R'_I = NR_I + 2M_3.$$

Let $\nu_B = \Pi_X(\mu_B)$ and $\nu_I = \Pi_X(\mu_I)$ be, respectively, the induced markings of $\mu_B$ and $\mu_I$ on $X$. We will regard $\nu_B$ as the base marking in $\mathrm{Mark}(X)$. Since $\mu_I \in \widetilde{\mathrm{Fix}}_{R_I}(f)$, for any $Z \subseteq X$,

$$
\begin{aligned}
d_Z(\nu_I, F\nu_I) &\leq d_Z(\mu_I, F\mu_I) + 2M_3 \\
&= d_Z(\mu_I, f^{L_X+1}\mu_I) + 2M_3 \\
&\leq \sum_{i=0}^{L_X} d_Z(f^i\mu_I, f^{i+1}\mu_I) + 2M_3 \\
&= (L_X + 1)\, d_Z(\mu_I, f\mu_I) + 2M_3 \\
&\leq (L_X + 1)R_I + 2M_3 \\
&< R'_I.
\end{aligned}
$$

In other words, $\nu_I \in \widetilde{\mathrm{Fix}}_{R'_I}(F)$. By Equation (2), we have

$$
\begin{aligned}
d_X(\mu_B, \mu_I) &\leq d_X(\mu_B, \nu_B) + d_X(\nu_B, \nu_I) + d_X(\nu_I, \mu_I) \\
&\leq d_X(\nu_B, \nu_I) + 2M_3.
\end{aligned}
$$

Using the above inequality and the fact that $X \in \Omega(\mu_I, R_I)$, we obtain

$$
\begin{aligned}
d_X(\nu_B, \nu_I) &\geq d_X(\mu_B, \mu_I) - 2M_3 \\
&> N \max_{0 \leq i \leq L_X} \{ d_{f^i(X)}(\mu_B, f\mu_B) \} + NR_I + \Theta - 2M_3 \\
&> \sum_{i=0}^{L_x} d_{f^{-i}(X)}(\mu_B, f\mu_B) + NR_I + \Theta - 2M_3 \\
&= \sum_{i=0}^{L_X} d_X(f^i\mu_B, f^{i+1}\mu_B) + NR_I + \Theta - 2M_3 \\
&\geq d_X(\mu_B, f^{L_X+1}\mu_B) + NR_I + \Theta - 2M_3 \\
&\geq d_X(\nu_B, F\nu_B) + NR_I + \Theta - 4M_3 \\
&\geq d_X(\nu_B, F\nu_B) + R'_I.
\end{aligned}
$$

In other words, $X \in \Omega(\nu_I, R'_I, F)$. We may apply Proposition 4.2.1, treating $X$ as the whole surface. This gives a marking $\nu_O$ on $X$ and a constant $R'_O \geq R'_I$ depending only on $R'_I$ (hence $R_I$) such that

(P1) For any $Z \subseteq X$,

$$d_Z(\nu_O, F\nu_O) \leq R'_O. \tag{30}$$

(P2) $X \notin \Omega(\nu_O, R'_O, F)$, meaning

$$d_X(\nu_B, \nu_O) \le d_X(\nu_B, F\nu_B) + R'_O. \tag{31}$$

The action of $f$ induces a marking $f^i \nu_O$ on each $f^i(X)$. We complete

$$c \cup \bigcup_{i=0}^{L_X} f^i \nu_O$$

to a marking $\mu_O$ on $\mathcal{S}$ by extending $\mu_I$ to the remaining complements and the curves in $c$. In this case, set $R_O = R'_O + 2M_3$.

Now for $X$ either a curve or a non-annular domain, let $\mu_O$ and $R_O$ be the appropriate marking and constant. We will show $\mu_O$ and $R_O$ satisfy properties (Q1), (Q3), (Q4) and (Q5). Let's consider the following analyses.

(c1)    If $Z = \mathcal{S}$, by assumption, $\mathcal{S} \notin \Omega(\mu_I, R_I)$, so

$$d_{\mathcal{S}}(\mu_B, \mu_I) \le d_{\mathcal{S}}(\mu_B, f\mu_B) + R_I.$$

Since $X$ is a domain of a geodesic in $H(\mu_B, \mu_I)$, we have

$$\begin{aligned}
d_{\mathcal{S}}(\mu_B, \mu_O) &\le d_{\mathcal{S}}(\mu_B, c) + 2 \\
&\le d_{\mathcal{S}}(\mu_B, \mu_I) + M_3 + 2 \\
&\le d_{\mathcal{S}}(\mu_B, f\mu_B) + R_I + M_3 + 2 \\
&\le d_{\mathcal{S}}(\mu_B, f\mu_B) + R_O.
\end{aligned}$$

In particular, $S \notin \Omega(\mu_O, R_O)$. As in (b1) (case II) of Proposition 4.2.1, we also have

$$d_{\mathcal{S}}(\mu_O, f\mu_O) \le 4.$$

(c2)    If $Z \ne \mathcal{S}$ but some curve of $c$ crosses $Z$, then the same argument of (b2) of Proposition 4.2.1 applies to give

$$d_Z(\mu_B, \mu_O) \le d_Z(\mu_B, \mu_I) + M_3 + 4,$$

and

$$d_Z(\mu_O, f\mu_O) \le 8.$$

(c3)    If $Z$ is a subsurface of some component domain of $(\mathcal{S}, c)$ on which $\mu_O$ is induced from $\mu_I$ (this includes the possibility that $Z$ is a curve in $c$ when $X$ is not a curve), then, as in (b3) of Proposition 4.2.1,

$$d_Z(\mu_B, \mu_O) \le d_Z(\mu_B, \mu_I) + M_3,$$

and

$$d_Z(\mu_O, f\mu_O) \le R_I + 2M_3 \le R_O.$$

(c4) If $X$ is a curve, then by construction

$$\begin{aligned} d_X(\mu_B, \mu_O) &\leq d_X(\mu_B, b) + M_3 \\ &\leq d_X(\mu_B, f\mu_B) + 2\delta + M_3 \\ &\leq d_X(\mu_B, f\mu_B) + R_O, \end{aligned}$$

and

$$\begin{aligned} d_X(\mu_O, f\mu_O) &\leq d_X(b, f(b)) + 2M_3 \\ &\leq 2\delta + 2M_3 \\ &\leq R_O. \end{aligned}$$

If $X$ is non-annular and $Z \subseteq X$, then it follows from (30) that

$$\begin{aligned} d_Z(\mu_O, f\mu_O) &\leq d_Z(\nu_O, F\nu_O) + 2M_3 \\ &\leq R'_O + 2M_3 \\ &= R_O. \end{aligned}$$

Finally, (31) yields

$$\begin{aligned} d_X(\mu_B, \mu_O) &\leq d_X(\nu_B, \nu_O) + 2M_3 \\ &\leq d_X(\nu_B, F\nu_B) + R'_O + 2M_3 \\ &\leq d_X(\mu_B, f^{L_X+1}\mu_B) + R'_O + 4M_3 \\ &\leq N \max_{0 \leq i \leq L_X} \{d_{f^i(X)}(\mu_B, f\mu_B)\} + NR_O + \Theta. \quad (32) \end{aligned}$$

One consequence here is that, whether or not $X$ is a curve, $X \notin \Omega(\mu_O, R_O)$.

(c5) If $X$ is a curve and $0 < i \leq L_X$, since both $\mu_O$ and $f\mu_O$ contain $f^i(b)$ (as a transversal), they are $M3$-close to $f^i(b)$ in $\mathcal{C}(f^i(X))$. Hence

$$d_{f^i(X)}(\mu_O, f\mu_O) \leq d_{f^i(X)}(f^i(b), f^i(b)) + 2M_3 e R_O.$$

Furthermore,

$$\begin{aligned} d_{f^i(X)}(\mu_B, \mu_O) &\leq d_{f^i(X)}(\mu_B, f^i(b)) + M_3 \\ &\leq d_X(f^{-i}\mu_B, b) + M_3 \\ &\leq d_X(\mu_B, b) + d_X(\mu_B, f^{-i}(\mu_B)) + M_3 \\ &\leq d_X(\mu_B, b) + \sum_{j=0}^{i-1} d_X(f^{-j}\mu_B, f^{-(j+1)}\mu_B) + M_3 \\ &= d_X(\mu_B, f\mu_B) + 2\delta + \sum_{j=0}^{i-1} d_{f^{j+1}(X)}(\mu_B, f\mu_B) + M_3 \\ &= \sum_{j=0}^{i} d_{f^j(X)}(\mu_B, f\mu_B) + 2\delta + M_3 \\ &\leq N \max_{0 \leq j \leq L_X} \{d_{f^j(X)}(\mu_B, f\mu_B)\} + R_O. \end{aligned}$$

If $X$ is non-annular and $Z \subseteq f^i(X)$, $0 < i \leq L_X$, then

$$d_Z(\mu_O, f\mu_O) = d_Z(f^i\nu_O, f^i\nu_O) + 2M_3 \leq R_O.$$

For $f^i(X)$, $0 < i \leq L_X$, then

$$\begin{aligned}
d_{f^i(X)}(\mu_B, \mu_O) &\leq d_{f^i(X)}(\mu_B, f^i\nu_O) + M_3 \\
&= d_X(f^{-i}\mu_B, \nu_O) + M_3 \\
&\leq d_X(\mu_B, \nu_O) + \sum_{j=0}^{i-1} d_X(f^{-j}\mu_B, f^{-(j+1)}\mu_B)
\end{aligned}$$

By (32) $\leq 2N \max_{0 \leq j \leq L_X} \{d_{f^j(X)}(\mu_B, f\mu_B)\} + NR_O + \Theta.$

An consequence here is that $f^i(X) \notin \Omega(\mu_O, R_O)$.

Together from (c1) to (c5), we have shown that $\mu_O \in \widetilde{\mathrm{Fix}}_{R_O}(f)$. Property (Q3) is verified in cases (c4) and (c5). To see (Q5), if $Z \in \Omega(\mu_O, R_O)$, then $Z$ is either of case (c2), (c3), or $Z \subsetneq f^i(X)$, for some $i = 0, \ldots, L_X$. In (c2) or (c3), since

$$d_Z(\mu_B, \mu_O) \leq d_Z(\mu_B, \mu_I) + M_3 + 4,$$

it follows that

$$Z \in \Omega(\mu_O, R_O) \implies Z \in \Omega(\mu_I, R_I).$$

To see (Q4), we use Lemma 2.8.2 on the assumption $X <_t f^{-i}(Z)$ to obtain

$$d_{f^{-i}(Z)}(\mu_B, \partial X) \leq M_1.$$

Since $\mu_O$ contains $c = \bigcup_j \partial f^j(X)$ as base curves, we have

$$\begin{aligned}
d_Z(\mu_B, \mu_O) &= d_{f^{-i}(Z)}(f^{-i}\mu_B, f^{-i}\mu_O) \\
&\leq d_{f^{-i}(Z)}(f^{-i}\mu_B, \mu_B) + d_{f^{-i}(Z)}(\mu_B, f^{-i}\mu_O) \\
&\leq N \max_{0 \leq i \leq L_Z} \{d_{f^i(Z)}(\mu_B, f\mu_B)\} + d_{f^{-i}(Z)}(\mu_B, \partial X) + M_3 \\
&\leq N \max_{0 \leq i \leq L_Z} \{d_{f^i(Z)}(\mu_B, f\mu_B)\} + M_1 + M_3.
\end{aligned}$$

Therefore, $Z \notin \Omega(\mu_O, R_O)$. This concludes (Q4) and the proof for $n = 1$.

For $n > 1$, the proof is essentially the same. By re-indexing if necessary, we may assume $X_1, \ldots X_m$, $m \leq n$, are non-annular domains, and $X_{m+1}, \ldots, X_n$ are curves. The assumption that $\mathcal{U}_1, \ldots, \mathcal{U}_n$ are pairwise disjoint allows us to apply the proof of case $n = 1$ to all $X_j$'s simultaneously. More precisely, for each $1 \leq j \leq m$, let $\nu_{O,j}$ be the marking on $X_j$ coming from case $n = 1$. Similarly, for $m + 1 \leq j \leq n$, let $b_j$ be the transversal curve to $X_j$ coming from case $n = 1$. Let $c = \cup_j c_j$ where $c_j = \partial X_j \cup \cdots \cup f^{L_{X_j}}(X_j)$. The set

$$\cup \left( \bigcup_{j=1}^{m} \bigcup_{i=0}^{L_{X_j}} f^i\nu_{O,j} \right) \cup \left( \bigcup_{i=1}^{L_{X_j}} \bigcup_{j=m+1}^{n} f^i(b_j) \right)$$

can be extended to a marking $\mu_O$ by extending $\mu_I$ to the remaining complements and the curves in $c_1, \ldots, c_m$. Let $R_O$ be the maximum of the two constants from case $n = 1$ (one constant for a curve and one constant for a non-annular domain). The proof that $\mu_O$ and $R_O$ satisfy the desired properties (Q1)–(Q5) is the same as the proof for $n = 1$.                                                                $\square$

### 4.3 Reducing complexity.

In this section, we show how to use Propositions 4.2.1 and 4.2.2 to construct $R$ and $\mu$ for Theorem 4.0.2.

From now on, a pair $(\mu, R)$ will always mean $\mu \in \widetilde{\mathrm{Fix}}_R(f)$.

PROPOSITION 4.3.1 (Reducing complexity). *Let $f \in \mathcal{MCG}(\mathcal{S})$ be of finite order. Let $R_I$ and $\mu_I$ be as in Proposition 4.2.1. Suppose $\Omega(\mu_I, R_I) \neq \emptyset$. There exists $R_O$, depending only on $R_I$, and $\mu_O \in \widetilde{\mathrm{Fix}}_{R_O}(f)$ such that*

$$\xi(\mu_O, R_O) < \xi(\mu_I, R_I).$$

*Proof.* If $\mathcal{S} \in \Omega(\mu_I, R_I)$, then Proposition 4.2.1 produces $(\mu_O, R_O)$ such that $R_O$ depends only on $R_I$ and $\mathcal{S} \notin \Omega(\mu_O, R_O)$, hence $\xi(\mu_O, R_O) < \xi(\mathcal{S}) = \xi(\mu_I, R_I)$.

Now suppose $\mathcal{S} \notin \Omega(\mu_I, R_I)$. Choose a maximal element $X_1 \in \Omega(\mu_I, R_I)$. This in particular means $X_1$ has maximal complexity over all elements of $\Omega(\mu_I, R_I)$. Set $\mathcal{U}_1 = X_1 \cup \cdots \cup f^{L_{X_1}}(X_1)$. Consider the maximal complexity of the elements in $\Omega(\mu_I, R_I)$ supported on $\mathcal{S} \setminus \mathcal{U}_1$. If this complexity is strictly less than $\xi(X_1)$, then we stop. If this complexity is not strictly less than $\xi(X_1)$, then we may choose $X_2$ of maximal order in $\Omega(\mu_I, R_I)$ supported on $\mathcal{S} \setminus \mathcal{U}_1$ such that $\xi(X_2) = \xi(X_1)$. Set $\mathcal{U}_2 = X_2 \cup \cdots \cup f^{L_{X_2}}(X_2)$. In this case, $\mathcal{U}_1$ and $\mathcal{U}_2$ are disjoint. Now we repeat this process by considering the maximum complexity of the elements in $\Omega(\mu_I, R_I)$ supported on $\mathcal{S} \setminus (\mathcal{U}_1 \cup \mathcal{U}_2)$. Continuing this way, we eventually *exhaust* $\mathcal{S}$ by a sequence

$$\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_n$$

in the following sense:

- For each $i$, the set $\mathcal{U}_i$ is a disjoint union of subsurfaces of $\mathcal{S}$ of the form

$$\mathcal{U}_i = X_i \cup \cdots \cup f^{L_{X_i}}(X_i),$$

  with $\xi(X_i) = \xi(X_1)$.
- The sets $\mathcal{U}_1, \ldots, \mathcal{U}_n$ are pairwise disjoint.
- The maximal complexity of the bad domains in $\Omega(\mu_I, R_I)$ supported on

$$\mathcal{S} \setminus (\mathcal{U}_1 \cup \cdots \cup \mathcal{U}_n)$$

  is strictly less than $\xi(X_1)$.

Note that the exhaustion sequence has length $n$ which is bounded uniformly by a constant depending only on $\mathcal{S}$. Denote by

$$c_i = \partial X_i \cup \cdots \cup \partial f^{L_{X_i}}(X_i).$$

By assumption, $\mathcal{U}_1, \ldots, \mathcal{U}_n$ are pairwise disjoint, so we can apply Proposition 4.2.2 to construct a pair $(\mu_O, R_O)$ with $\mu_O$ containing $c_1 \cup \cdots \cup c_n$ as base curves and $R_O$ depending only on $R_I$. By properties (Q3), (Q4), and (Q5) of Proposition 4.2.2, if $Z \in \Omega(\mu_O, R_O)$, then either

(i) $Z$ intersects some curve in $c_i$ for some $i$ (see case (c2) in the proof of Proposition 4.2.2).

(ii) $Z$ is supported on $\mathcal{S} \setminus (\mathcal{U}_1 \cup \cdots \cup \mathcal{U}_n)$. (See case (c3) in the proof of Proposition 4.2.2.)

(iii) $Z \subsetneq f^j(X_i)$, for some $0 \leq i \leq n$ and $0 \leq j \leq L_{X_i}$.

Immediately, case (iii) has $\xi(Z) < \xi(X_1)$. Recall that for either case (i) or (ii),

$$Z \in \Omega(\mu_O, R_O) \implies Z \in \Omega(\mu_I, R_I).$$

Since $\mathcal{S}$ is exhausted by assumption, case (ii) also means $\xi(Z) < \xi(X_1)$. Lastly, suppose $Z$ is of case (i). Choose the minimal index $i$ such that $Z$ intersects a curve in $c_i$. In other words, $Z$ is supported on

$$\mathcal{S} \setminus (\mathcal{U}_1 \cup \cdots \cup \mathcal{U}_{i-1}),$$

and $Z$ interlocks $f^j(X_i)$, for some $j$. Our choice of $X_i$ has maximal order among the bad domains supported on

$$\mathcal{S} \setminus (\mathcal{U}_1 \cup \cdots \cup \mathcal{U}_{i-1}).$$

Therefore, if $\xi(Z) = \xi(X_1) = \xi(X_i)$, then $X_i <_t f^{-j}(Z)$ in $H(\mu_B, \mu_I)$. Property (Q5) of Proposition 4.2.2 guarantees that such domains do not appear in $\Omega(\mu_O, R_O)$. Thus, any $Z$ of case (i) must also have $\xi(Z) < \xi(X)$. $\square$

Let $R_1$ be the minimal constant satisfying Lemma 3.1.4.

COROLLARY 4.3.2 (Termination). *There exists $R \geq R_1$ depending only on $\mathcal{S}$ such that any finite order $f \in \mathcal{MCG}(\mathcal{S})$ has a $\mu \in \widetilde{\mathrm{Fix}}_R(f)$ satisfying $\Omega(\mu, R) = \emptyset$.*

*Proof.* Let $\mu_1 \in \widetilde{\mathrm{Fix}}_{R_1}(f)$. If $\Omega(\mu_1, R_1) = \emptyset$, then we are done. If not, then applying Proposition 4.3.1 iteratively yields a sequence of pairs $(\mu_1, R_1), (\mu_2, R_2), \ldots$ such that

- $\mu_{i+1} \in \widetilde{\mathrm{Fix}}_{R_{i+1}}(f)$, where $R_{i+1}$ depends only on $R_i$.
- $\xi_{i+1} \lneq \xi_i$, where $\xi_i = \xi(\mu_i, R_i)$.

Since $\xi_i$ corresponds to the maximum complexity over elements in $\Omega(\mu_i, R_i)$, and that $\xi_i$'s are strictly decreasing, we must eventually reach a pair $(\mu_n, R_n)$ for which $\xi_n = -2$, i.e. $\Omega(\mu_n, R_n) = \emptyset$. Moreover, since elements of $\Omega(\mu_i, R_i)$ come from subsurfaces of $\mathcal{S}$, $\xi_1 \leq \xi(\mathcal{S}) = 3g - 3 + b$. This gives a bound on $n \leq 3g - 1 + b$, and therefore $R_n$ depends only on $\mathcal{S}$. $\square$

This concludes the proof of Theorem 4.0.2.

## 5  L.B.C. Property for Reducible Mapping Classes

In this section, we prove L.B.C. property for (infinite-order) reducible elements of $\mathcal{MCG}(\mathcal{S})$. We would like to use an induction argument on subsurfaces. To do so, we make the following observation.

Let $f \in \mathcal{MCG}(\mathcal{S})$ be reducible with canonical reducing system $\sigma$ (see Section 2.2.3). If $\mu \in \mathrm{Mark}(\mathcal{S})$ is a marking containing $\sigma$ as base curves, then so does $f\mu$. This means that any hierarchy $H(\mu, f\mu)$ decomposes into geodesics supported on component domains of $(\mathcal{S}, \sigma)$. For such a marking $\mu$, we can control each component domain independently, allowing the arguments of Theorem 2.2.4 and Corollary 4.0.4 to pass through to subsurfaces. This inspires the definition of a good marking for $f$ (Definition 5.0.3). We construct a finite collection of good markings and prove Theorem E of the introduction. The induction argument on subsurfaces using good markings appears in Proposition 5.0.5. Finally, Corollary 5.0.6 combines the finiteness and the induction argument to finish the proof of L.B.C. property for reducible elements of $\mathcal{MCG}(\mathcal{S})$.

DEFINITION 5.0.3 (Good marking). *Let $f \in \mathcal{MCG}(\mathcal{S})$ be an infinite-order reducible element and let $\sigma$ be its associated canonical reducing system. We say a marking $\mu \in \mathrm{Mark}(\mathcal{S})$ is a good marking for $f$ if $\sigma \subseteq \mathrm{base}(\mu)$.*

Up to homeomorphisms of $\mathcal{S}$, there are only finitely many multicurves on $\mathcal{S}$. If $f \in \mathcal{MCG}(\mathcal{S})$ is reducible and $f = \omega^{-1}g\omega$, then $\omega(\sigma_f) = \sigma_g$, where $\sigma_f$ and $\sigma_g$ are canonical reducing system for $f$ and $g$, respectively. We fix a representative for each homeomorphism type of a multicurve. Further, for each representative multicurve $\sigma$, we complete $\sigma$ into a finite set of markings representing each homeomorphism type of a marking containing $\sigma$ as base curves. Let $\mathcal{M}$ the collection of all such *representative markings*, one from each homeomorphism type. The following is Theorem E of the introduction.

**Theorem 5.0.4.**  *There exists $a \in \mathcal{MCG}(\mathcal{S})$ such that $a^{-1}fa$ has a good marking in $\mathcal{M}$, and*

$$d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, a\mu_B) \prec d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f\mu_B).$$

*Furthermore, if $f$ and $g$ are conjugate and $b^{-1}gb = a^{-1}fa$ with $d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, b\mu_B) \prec d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, g\mu_B)$, then we may choose the same good marking in $\mathcal{M}$ for $b^{-1}gb$.*

Before proving Theorem 5.0.4, we set up some notations. Fix a multicurve $\sigma$ on $\mathcal{S}$. Let $\mathrm{Stab}(\sigma) < \mathcal{MCG}(\mathcal{S})$ be the subgroup stabilizing $\sigma$ as a set:

$$\mathrm{Stab}(\sigma) : \{h \in \mathcal{MCG}(\mathcal{S}) : h(\sigma) = \sigma\}.$$

The action of $\mathrm{Stab}(\sigma)$ on the complementary components of $\mathcal{S} \setminus \sigma$ induces an exact sequence:

$$1 \longrightarrow \mathrm{Stab}_0(\sigma) \longrightarrow \mathrm{Stab}(\sigma) \xrightarrow{\pi} \text{Finite Group} \longrightarrow 1. \tag{33}$$

Consider the kernel $\text{Stab}_0(\sigma)$ of the above sequence. If $f \in \text{Stab}_0(\sigma)$, then $f$ acts on each complementary component $Y \subset \mathcal{S} \setminus \sigma$. In other words, $f$ defines an element $f|_Y \in \mathcal{MCG}_0(Y)$ for each $Y \subset \mathcal{S} \setminus \sigma$, where $\mathcal{MCG}_0(Y) \subset \mathcal{MCG}(Y)$ is the subgroup fixing $\partial Y$. We have an exact sequence:

$$1 \longrightarrow T_\sigma \longrightarrow \text{Stab}_0(\sigma) \longrightarrow \prod_{Y \in \mathcal{S} \setminus \sigma} \mathcal{MCG}_0(Y) \longrightarrow 1, \qquad (34)$$

where $T_\sigma$ is a free abelian group with basis the Dehn twists along curves in $\sigma$. By the classification theorem, $f|_Y$ is either pseudo-Anosov or has finite order. We say an element $f \in \text{Stab}_0(\sigma)$ is *pure* if $f|_Y$ is either pseudo-Anosov or the identity on $Y$. The order of the finite group in (33) is bounded by a constant $N$ depending only on $\mathcal{S}$. Thus, for any $\sigma$ and any $f \in \text{Stab}(\sigma), f^N \in \text{Stab}_0(\sigma)$. Moreover, since there are only finitely many subsurfaces of $\mathcal{S}$ up to homeomorphism, one can choose a constant for Corollary 3.1.2 that works for $\mathcal{S}$ and all subsurfaces of $\mathcal{S}$. Thus, there exists some universal power $N = N(S)$ depending only $\mathcal{S}$, such that for any reducible mapping element $f \in \mathcal{MCG}(\mathcal{S}), f^N$ is pure.

We can characterize the canonical reducing system for a reducible mapping class as follows. Suppose $f \in \text{Stab}(\sigma)$. Let $g = f^N$ be pure. Then $\sigma = \sigma_f$ is the canonical reducing system for $f$ if for any $\alpha \in \sigma$, one of the following holds.

(H1) There exists a domain $Y$ in $\mathcal{S} \setminus \sigma$ such that $\alpha$ is a boundary component of $Y$ and $g|_Y$ is pseudo-Anosov on $Y$.

(H2) There exists a domain $Z \subset \mathcal{S}$ such that $g|_Z$ is a non-zero power of a Dehn twist along $\alpha$.

To see this, let $\alpha \in \sigma$ be such that condition (H1) does not hold. Then $\alpha$ must bound two (not necessarily distinct) components $X$ and $Y$ of $\mathcal{S} \setminus \sigma$ such that $g|_X$ and $g|_Y$ are both the identity. In this case, let $Z = X \cup Y \cup \alpha$. Then $g|_Z$ is a non-zero power of Dehn twist along $\alpha$. Otherwise, the first return map of $f$ to $Z$ is of finite order, and one can thus obtain a smaller reducing system for $f$ by removing $\alpha$, contradicting minimality of $\sigma$. Note that this proof also implies that if $\sigma$ is a canonical reducing system for $f$, then $\sigma$ is also the canonical reducing system for any power of $f$.

In case (H2), it follows that for any $n \in \mathbb{N}$ and any $v \in \mathcal{C}(\alpha)$,

$$d_\alpha(v, g^n(v)) \geq |n|.$$

Compare this with Lemma 2.3.2. Since there are only finitely many domains of $\mathcal{S}$ up to homeomorphism, we can choose $N_0$ depending only on $\mathcal{S}$ such that the following holds. Let $M_2$ be the constant of Lemma 2.6.3. For any multicurve $\sigma$ and any $g \in \text{Stab}_0(\sigma)$, let $Y$ be either a component of $\mathcal{S} \setminus \sigma$ on which $g$ is pseudo-Anosov, or $Y$ is a curve in $\sigma$ such that property (H2) holds, then for any $n \geq N_0$ and any $v \in \mathcal{C}(Y)$,

$$d_Y(v, g^n(v)) \geq M_2.$$

For any $g \in \mathrm{Stab}_0(\sigma)$, Equation (3) has the following consequence. If $\mu \in \mathrm{Mark}(\mathcal{S})$ is a good marking for $g$, then

$$d_{\mathrm{Mark}(\mathcal{S})}(\mu, g\mu) \asymp \sum_{X \subset \mathcal{S}\backslash\sigma} d_{\mathrm{Mark}(X)}(\mu, g|_X\mu) + \sum_{\alpha \in \sigma} d_\alpha(\mu, g\mu), \tag{35}$$

where

$$d_{\mathrm{Mark}(X)}(\mu, g|_X\mu) := d_{\mathrm{Mark}(X)}(\Pi_X(\mu), g|_X\Pi_X(\mu)) \asymp d_{\mathrm{Mark}(X)}(\Pi_X(\mu), \Pi_X(g\mu)).$$

We will say an element $g \in \mathrm{Stab}_0(\sigma)$ *does not twist* along $\alpha \in \sigma$ if for any $v \in \mathcal{C}(\alpha)$,

$$\lim_{n\to\infty} \frac{d_\alpha(v, g^n(v))}{n} = 0.$$

In this case, $d_\alpha(\mu, g\mu) \prec 1$, and one can ignore the second summand on the right hand side of Equation (35).

*Proof of Theorem* 5.0.4. The set $\mathcal{M}$ is finite and each conjugacy class of $\mathcal{MCG}(\mathcal{S})$ has a good marking in $\mathcal{M}$ by construction. Let

$$C = \max\{d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \mu) : \mu \in \mathcal{M}\}.$$

Let $f \in \mathcal{MCG}(\mathcal{S})$ be reducible with canonical reducing system $\sigma_f$ and set $F = f^{N_0}$. Let $\mu \in \mathcal{M}$ be arbitrary. By our definition of $N_0$, for any $\alpha \in \sigma_f, \alpha$ is either a boundary curve of a domain $Y$ such that $d_Y(\mu, F\mu) \geq M_2$, or $d_\alpha(\mu, F\mu) \geq M_2$. Thus, by Lemma 2.6.3, $\alpha$ is either a domain for a geodesic in $H(\mu, F\mu)$ or is a boundary curve of a domain for a geodesic in $H(\mu, F\mu)$. A consequence is that, for any $Y \subset \mathcal{S}$ that intersects a curve $\alpha \in \sigma_f$, by choosing a hierarchal slice containing $\alpha$ and using Lemma 2.7.2, we have

$$d_Y(\alpha, F\mu) \prec d_Y(\mu, F\mu). \tag{36}$$

Now let $\mu'$ be a marking extension of $\sigma_f$ relative to $F\mu$. By Lemma 2.5.7, for any component domain $Y$ of $(\mathcal{S}, \sigma_f), d_Y(\mu', F\mu)$ is uniformly bounded. On the other hand, if $Y \subseteq \mathcal{S}$ is any domain that intersects some curve $\alpha \in \sigma_f$, then by (36) and the fact that $\sigma_f \subseteq \mathrm{base}(\mu')$, we have

$$d_Y(\mu', F\mu) \leq d_Y(\alpha, F\mu) + 2 \prec d_Y(\mu, F\mu).$$

By ranging over all $Y \subseteq \mathcal{S}$ on which $d_Y(\mu', F\mu)$ is sufficiently large, we obtain

$$d_{\mathrm{Mark}(\mathcal{S})}(\mu', F\mu) \prec d_{\mathrm{Mark}(\mathcal{S})}(\mu, F\mu).$$

This implies:

$$\begin{aligned}
d_{\mathrm{Mark}(\mathcal{S})}(\mu, \mu') &\leq d_{\mathrm{Mark}(\mathcal{S})}(\mu, F\mu) + d_{\mathrm{Mark}(\mathcal{S})}(F\mu, \mu') \\
&\prec d_{\mathrm{Mark}(\mathcal{S})}(\mu, F\mu) \\
&\prec d_{\mathrm{Mark}(\mathcal{S})}(\mu, f\mu).
\end{aligned}$$

Let $\sigma$ be the representative multicurve for $\sigma_f$ and let $\mathcal{M}(\sigma) \subset \mathcal{M}$ be the subset of markings containing $\sigma$ as base curves. Now choose a marking $\mu'' \in \mathcal{M}(\sigma)$ such that there exist $a \in \mathcal{MCG}(\mathcal{S})$ with $a(\mu'') = \mu'$. By construction, $a^{-1}fa$ has canonical reducing system $a^{-1}(\sigma_f) = \sigma \subseteq \text{base}(\mu'')$. We also have

$$
\begin{aligned}
d_{\text{Mark}(\mathcal{S})}(\mu_B, a\mu_B) &\leq d_{\text{Mark}(\mathcal{S})}(\mu_B, a\mu'') + d_{\text{Mark}(\mathcal{S})}(a\mu'', a\mu_B) \\
&\leq d_{\text{Mark}(\mathcal{S})}(\mu_B, \mu') + C \\
&\leq d_{\text{Mark}(\mathcal{S})}(\mu_B, \mu) + d_{\text{Mark}(\mathcal{S})}(\mu, \mu') + C \\
&\prec d_{\text{Mark}(\mathcal{S})}(\mu, f\mu) + 2C \\
&\leq d_{\text{Mark}(\mathcal{S})}(\mu, \mu_B) + d_{\text{Mark}(\mathcal{S})}(\mu_B, f\mu_B) + d_{\text{Mark}(\mathcal{S})}(f\mu_B, f\mu) + 2C \\
&\leq d_{\text{Mark}(\mathcal{S})}(\mu_B, f\mu_B) + 4C.
\end{aligned}
$$

If $g \in \mathcal{MCG}(\mathcal{S})$ is conjugate to $f$, then $\sigma$ would also be the representative multicurve for $\sigma_g$. Our construction produces an element $b \in \mathcal{MCG}(\mathcal{S})$ such that $b^{-1}gb$ has a good marking in $\mathcal{M}(\sigma)$ and $d_{\text{Mark}(\mathcal{S})}(\mu_B, b\mu_B) \prec d_{\text{Mark}(\mathcal{S})}(\mu_B, g\mu_B)$. Since any marking in $\mathcal{M}(\sigma)$ is a good marking for $b^{-1}gb$, including $\mu''$, and $\mathcal{M}(\sigma)$ is a finite set, the second statement follows. $\qquad \square$

PROPOSITION 5.0.5. *Suppose $f, g \in \mathcal{MCG}(\mathcal{S})$ are two conjugate infinite-order reducible mapping classes with the same canonical reducing system $\sigma$. Let $\mu$ be a good marking for $f$ and $g$. Then there exist a constant $K_\mu$ and $\omega \in \mathcal{MCG}(\mathcal{S})$ such that $\omega$ is a conjugator for $f$ and $g$, and*

$$
d_{\text{Mark}(\mathcal{S})}(\mu, \omega\mu) \prec K_\mu (d_{\text{Mark}(\mathcal{S})}(\mu, f\mu) + d_{\text{Mark}(\mathcal{S})}(\mu, g\mu)).
$$

*Proof.* Elements of $\text{Stab}_0(\sigma)$ are easier to handle, but a conjugator for $f^n$ and $g^n$ is not a conjugator for $f$ and $g$. Thus we cannot apply the results of Corollary 4.0.4 and Theorem 2.2.4, to powers of $f$ and $g$. We must deal with the issue of permuting subsurfaces in the proof of Proposition 5.0.5. Fix a finite collection $\mathcal{P} \subset \text{Stab}(\sigma)$ such that $\pi(\mathcal{P})$ in the exact sequence (33) is onto. Let

$$
P = \max\{d_{\text{Mark}(\mathcal{S})}(\mu, a\mu) : a \in \mathcal{P}\}.
$$

Choose $a \in \mathcal{P}$ such that $f$ and $g' = aga^{-1}$ have the following properties:

(i) $\pi(f) = \pi(g')$.
(ii) For any $X$ in $\mathcal{S} \setminus \sigma$, the first return map to $X$ of $f$ and $g'$ are conjugate.

If the proposition holds for $f$ and $g'$, say, there exist $K_\mu$ and $\omega \in \mathcal{MCG}(\mathcal{S})$ such that $f\omega = \omega g'$ and

$$
d_{\text{Mark}(\mathcal{S})}(\mu, \omega'\mu) \prec K_\mu (d_{\text{Mark}(\mathcal{S})}(\mu, f\mu) + d_{\text{Mark}(\mathcal{S})}(\mu, g'\mu)),
$$

then $\omega a$ and $K_\mu + 2PK_\mu + P$ verify the proposition for $f$ and $g$. Clearly, $\omega a$ is a conjugator for $f$ and $g$. It remains to check:

$$
\begin{aligned}
d_{\mathrm{Mark}(\mathcal{S})}(\mu, \omega a \mu) &\leq d_{\mathrm{Mark}(\mathcal{S})}(\mu, \omega\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\omega\mu, \omega a \mu) \\
&\leq d_{\mathrm{Mark}(\mathcal{S})}(\mu, \omega\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, a\mu) \\
&\prec K_\mu(d_{\mathrm{Mark}(\mathcal{S})}(\mu, f\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, aga^{-1}\mu)) + P \\
&\leq K_\mu(d_{\mathrm{Mark}(\mathcal{S})}(\mu, f\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, g\mu) + 2P) + P.
\end{aligned}
$$

Thus, we may assume $f$ and $g$ already satisfy properties (i) and (ii) above. We will find a conjugator $\omega \in \mathrm{Stab}_0(\sigma)$ for $f$ and $g$. To do this, we will use properties (i) and (ii) and the induction hypothesis to build a conjugating element $\omega_Y \in \mathcal{MCG}_0(Y)$ for each component $Y$ in $\mathcal{S} \setminus \sigma$. Then we will choose an appropriate lift $\omega \in \mathrm{Stab}_0(Y)$.

Decompose the complementary components of $\mathcal{S} \setminus \sigma$ into orbits under the action $f$. Pick a representative from each orbit. Let $X_1$ be one such representative and consider the sequence of distinct complementary subsurfaces of $\mathcal{S} \setminus \sigma$

$$
X_1, X_2, \ldots, X_n
$$

such that $f(X_i) = X_{i+1}$ and $g(X_i) = X_{i+1}$, for $i = 1, \ldots, n$ and $X_{n+1} = X_1$. Note that $n < N_0$. Set $f_i = f|_{X_i} : X_i \to X_{i+1}$, and similarly for $g_i$. Set the first return maps $F = f^{n+1}|_{X_1} \in \mathcal{MCG}(X_1)$ and $G = g^{n+1}|_{X_1} \in \mathcal{MCG}(X_1)$. The assumption is that $F$ and $G$ are conjugate in $\mathcal{MCG}(X_1)$. By Theorem 2.2.2, $F$ and $G$ are either pseudo-Anosov or have finite order on $X_1$. Letting $\nu_1 = \Pi_{X_1}(\mu)$ be the induced marking on $X_1$, it follows from results of Corollary 4.0.4 and Theorem 2.2.4 that there exist $\omega_1 \in \mathcal{MCG}_0(X_1)$ and $K_1 = K_1(\nu_1, X_1)$ such that $F\omega_1 = \omega_1 G$, and

$$
\begin{aligned}
d_{\mathrm{Mark}(X_1)}(\nu_1, \omega_1\nu_1) &\leq K_1(d_{\mathrm{Mark}(X_1)}(\nu_1, F\nu_1) + d_{\mathrm{Mark}(X_1)}(\nu_1, G\nu_1)) \\
&\prec K_1(d_{\mathrm{Mark}(\mathcal{S})}(\mu, F\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, G\mu)) \quad (\text{By } (3)) \\
&\prec K_1(d_{\mathrm{Mark}(\mathcal{S})}(\mu, f^n\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, g^n\mu)) \\
&\prec K_1(d_{\mathrm{Mark}(\mathcal{S})}(\mu, f\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, g\mu)).
\end{aligned}
$$

Using $f$ and $g$, we construct for each $X_i$ an element $\omega_i \in \mathcal{MCG}_0(X_i)$ such that $f_i\omega_i = \omega_{i+1}g_i$, for $i = 1, \ldots n$ and $n + 1 = 1$. The element $\omega_1 \in \mathcal{MCG}_0(X_1)$ is defined. For each $i = 1, \ldots n$, set

$$
\omega_{i+1} = f_i \cdots f_1 \omega_1 g_1^{-1} \cdots g_i^{-1}.
$$

In particular, $\omega_{n+1} = F\omega_1 G^{-1} = \omega_1$. Let $\nu_i = \Pi_{X_i}(\mu)$. We have

$$
\begin{aligned}
d_{\mathrm{Mark}(X_i)}(\nu_i, \omega_{i+1}\nu_i) &= d_{\mathrm{Mark}(X_i)}(\nu_i, f_i \cdots f_1 \omega_1 g_1^{-1} \cdots g_i^{-1}\nu_i) \\
&= d_{\mathrm{Mark}(X_i)}(\nu_i, f^i|_{X_1}\omega_1 g^{-i}|_{X_i}\nu_i) \\
&= d_{\mathrm{Mark}(X_1)}(f^{-i}|_{X_i}\nu_i, \omega_1 g^{-i}|_{X_i}\nu_i) \\
&\leq d_{\mathrm{Mark}(X_1)}(f^{-i}|_{X_i}\nu_i, \nu_1) + d_{\mathrm{Mark}(X_1)}(\nu_1, \omega_1 g^{-i}|_{X_i}\nu_i) \\
&= d_{\mathrm{Mark}(X_1)}(f^{-i}|_{X_i}\nu_i, \nu_1) + d_{\mathrm{Mark}(X_1)}(\omega_1^{-1}\nu_1, g^{-i}|_{X_i}\nu_i) \\
&\leq d_{\mathrm{Mark}(X_1)}(f^{-i}|_{X_i}\nu_i, \nu_1) + d_{\mathrm{Mark}(X_1)}(\omega_1^{-1}\nu_1, \nu_1) \\
&\quad + d_{\mathrm{Mark}(X_1)}(\nu_1, g^{-i}\nu_i) \\
&\prec d_{\mathrm{Mark}(\mathcal{S})}(\mu, f^i\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, g^i\mu) + d_{\mathrm{Mark}(X_1)}(\nu_1, \omega_1\nu_1) \\
&\prec K_1(d_{\mathrm{Mark}(\mathcal{S})}(\mu, f\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, g\mu)).
\end{aligned}
$$

We do this for each orbit of a complementary subsurface in $\mathcal{S} \setminus \sigma$, building for each $Y \subset \mathcal{S} \setminus \sigma$ an element $\omega_Y \in \mathcal{MCG}_0(Y)$. Consider any element $\omega \in \mathrm{Stab}_0(\sigma)$ such that $\omega|_Y = \omega_Y$. Since twistings commute, any $\omega$ will satisfy $f\omega = \omega g$ by construction. Thus, we can choose a lift $\omega$ that does not twist along any curves in $\sigma$. Let $\{K_i\}$ be the constants associated to each orbit and let $K_\mu = \max\{K_i\}$. Using previous work and Eq. (35), we obtain

$$
\begin{aligned}
d_{\mathrm{Mark}(\mathcal{S})}(\mu, \omega\mu) &\asymp \sum_{X \subset \mathcal{S}\setminus\sigma} d_{\mathrm{Mark}(X)}(\mu, \omega_X\mu) + \sum_{\alpha \in \sigma} d_\alpha(\mu, \omega\mu) \\
&\prec \sum_{X \subset \mathcal{S}\setminus\sigma} K_\mu(d_{\mathrm{Mark}(\mathcal{S})}(\mu, f\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, g\mu)) \\
&\prec K_\mu(d_{\mathrm{Mark}(\mathcal{S})}(\mu, f\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, g\mu)). \qquad \square
\end{aligned}
$$

COROLLARY 5.0.6 (L.B.C. property for reducible mapping classes). *If $f, g \in \mathcal{MCG}(\mathcal{S})$ are conjugate reducible mapping classes of infinite order, then there is a conjugating element $\omega \in \mathcal{MCG}(\mathcal{S})$ with*

$$
|\omega| \prec |f| + |g|.
$$

*Proof.* Let $\mathcal{M}$ be the set of representative markings and let $C$ be the constant bounding the diameter of $\mathcal{M}$. Let $K$ be the constant depending only on $\mathcal{S}$ defined by

$$
K = \max\{K_\mu : \mu \in \mathcal{M}\},
$$

where $K_\mu$ is the constant associated to $\mu$ from Proposition 5.0.5. Suppose $f_1$ and $f_2$ are conjugate reducible mapping classes of infinite order. Let $a_1$ and $a_2$ be such that $a_i f_i a_i^{-1} = g_i$ have a good marking $\mu \in \mathcal{M}$, and satisfying $d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, a_i\mu_B) \prec d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f_i\mu_B)$. Then each

$$
\begin{aligned}
d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, g_i\mu_B) &\leq 2d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, a_i\mu_B) + d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f_i\mu_B) \\
&\prec d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f_i\mu_B).
\end{aligned}
$$

By Proposition 5.0.5, there exists $\omega \in \mathcal{MCG}(\mathcal{S})$ such that $g_1\omega = \omega g_2$, and

$$d_{\mathrm{Mark}(\mathcal{S})}(\mu, \omega\mu) \prec K_\mu(d_{\mathrm{Mark}(\mathcal{S})}(\mu, g_1\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, g_2\mu))$$
$$\leq K(d_{\mathrm{Mark}(\mathcal{S})}(\mu, g_1\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, g_2\mu)).$$

Hence, by the triangle inequality,

$$d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \omega\mu_B) \leq d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, \omega\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\omega\mu, \omega\mu_B)$$
$$\leq 2C + Kd_{\mathrm{Mark}(\mathcal{S})}(\mu, g_1\mu) + d_{\mathrm{Mark}(\mathcal{S})}(\mu, g_2\mu))$$
$$\leq 2C + K(4C + d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, g_1\mu_B) + d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, g_2\mu_B))$$
$$\prec d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, g_1\mu_B) + d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, g_2\mu_B).$$

Set $\omega' = a_1^{-1}\omega a_2$. Then $\omega'$ is a conjugator for $f_1$ and $f_2$, and

$$d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \omega'\mu_B) \leq d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, a_1\mu_B) + d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, \omega\mu_B) + d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, a_2\mu_B)$$
$$\prec d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f_1\mu_B) + d_{\mathrm{Mark}(\mathcal{S})}(\mu_B, f_2\mu_B).$$

This concludes the proof the corollary. $\qquad\square$

## Acknowledgments

## References

[Ago11]   I. Agol. *Ideal Triangulations of Pseudo-Anosov Mapping Tori, Topology and Geometry in Dimension Three.* Contemp. Math., Vol. 560. Am. Math. Soc., Providence, RI (2011), pp. 1–17.

[BH95]    M. Bestvina and M. Handel. Train-tracks for surface homeomorphisms. *Topology*, (1)34 (1995), 109–140.

[BH99]    M.R. Bridson and A. Haefliger. *Metric Spaces of Non-Positive Curvature*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Vol. 319. Springer, Berlin (1999).

[BKMM12] J. Behrstock, B. Kleiner, Y. Minsky, and L. Mosher. Geometry and rigidity of mapping class groups. *Geometric Topology*, (2)16 (2012), 781–888.

[BM08]    J. Brock and H. Masur. Coarse and synthetic Weil–Petersson geometry: Quasi-flats, geodesics and relative hyperbolicity. *Geometric Topology*, (4)12 (2008), 2453–2495.

[Boo59]   W.W. Boone. The word problem. *Annals of Mathematics*, (2)70 (1959), 207–265.

[Bow06]    B.H. BOWDITCH. Intersection numbers and the hyperbolicity of the curve complex. *Journal für die Reine und Angewandte Mathematik*, 598 (2006), 105–129.

[Deh11]    M. DEHN. Über unendliche diskontinuierliche Gruppen. *Mathematische Annalen*, (1)71 (1911), 116–144.

[ECH+92]   D.B.A. EPSTEIN, J.W. CANNON, D.F. HOLT, S.V.F. LEVY, M.S. PATERSON, and W.P. THURSTON. *Word Processing in Groups*. Jones and Bartlett Publishers, Boston (1992).

[Far06]    B. FARB. *Some Problems on Mapping Class Groups and Moduli Space*, Problems on mapping class groups and related topics, Proc. Sympos. Pure Math., Vol. 74. Am. Math. Soc., Providence, RI (2006), pp. 11–55.

[FLP79]    *Travaux de Thurston sur les surfaces*, Astérisque, Vol. 66, Société Mathématique de France, Paris (1979), Séminaire Orsay, With an English summary.

[FM12]     B. FARB and D. MARGALIT. *A primer on Mapping Class Groups*, Princeton Mathematical Series, Vol. 49. Princeton University Press, Princeton (2012).

[Gro87]    M. GROMOV. *Hyperbolic Groups*, Essays in group theory, Mathematical Sciences Research Institute Publications, Vol. 8. Springer, New York (1987), pp. 75–263.

[Gro75]    E.K. GROSSMAN. On the residual finiteness of certain mapping class groups. *Journal of the London Mathematical Society*, (2)9 (1974/1975), 160–164.

[Ham09]    U. HAMENSTÄDT. *Geometry of the mapping class group, II: A biautomatic structure*, Preprint, 2009, Available at arXiv:math/0912.0137v1 [math.GR].

[Har81]    W.J. HARVEY. *Boundary Structure of the Modular Group*, Riemann surfaces and related topics: Proceedings of the 1978 Stony Brook Conference (State Univ. New York, Stony Brook, N.Y., 1978), Annals of Math. Stud., Vol. 97. Princeton Univ. Press, Princeton, N.J. (1981), pp. 245–251.

[Hem79]    G. HEMION. On the classification of homeomorphisms of 2-manifolds and the classification of 3-manifolds. *Acta Mathematica*, (1–2)142 (1979), 123–155.

[HT80]     A. HATCHER and W. THURSTON. A presentation for the mapping class group of a closed orientable surface. *Topology*, (3)19 (1980), 221–237.

[Ker83]    S.P. KERCKHOFF. The Nielsen realization problem. *The Annals of Mathematics*, (2)117 (1983), 235–265.

[Lic64]    W.B.R. LICKORISH. A finite set of generators for the homeotopy group of a 2-manifold. *Proceedings of the Cambridge Philosophical Society Journal*, 60 (1964), 769–778.

[Lys89]    I.G. LYSËNOK. Some algorithmic properties of hyperbolic groups. *Izvestiya Akademii Nauk SSSR Seriya Matematicheskaya*, (4)53 (1989), 814–832, 912.

[Man05]    J.F. MANNING. Geometry of pseudocharacters. *Geometric Topology*, 9 (2005), 1147–1185 (electronic).

[Min96]    Y.N. MINSKY. *A Geometric Approach to the Complex of Curves on a Surface*, Topology and Teichmüller spaces (Katinkulta, 1995). World Sci. Publ., River Edge, NJ (1996), pp. 149–158.

[Min10]    Y. MINSKY. The classification of Kleinian surface groups. I. Models and bounds. *Annals of Mathematics* (2), (1)171 (2010), 1–107.

[MM99]     H.A. MASUR and Y.N. MINSKY. Geometry of the complex of curves. I. Hyperbolicity. *Inventiones Mathematicae*, (1)138 (1999), 103–149.

[MM00]     H.A. MASUR and Y.N. MINSKY. Geometry of the complex of curves. II. Hierarchical structure. *Geomertic and Functional Analysis*, (4)10 (2000), 902–974.

[Mos86]    L. MOSHER. *The Classification of Pseudo-Anosovs*, Low-dimensional topology and Kleinian groups (Coventry/Durham, 1984), London Math. Soc. Lecture

Note Ser., Vol. 112. Cambridge University Press, Cambridge (1986), pp. 13–75.

[Mos95]   L. MOSHER. Mapping class groups are automatic. *Annals of Mathematics* (2), (2)142 (1995), 303–384.

[Mos03]   L. MOSHER. *Train track expansions of measured foliations*, Preprint, 2003, Available at [andromeda.rutgers.edu/~mosher](andromeda.rutgers.edu/~mosher).

[Mos07]   L. MOSHER. *MSRI Course on mapping class groups*, Lecture notes, 2007, Available at [andromeda.rutger.edu/~mosher](andromeda.rutger.edu/~mosher).

[Nov58]   P.S. NOVIKOV. *On the Algorithmic Insolvability of the Word Problem in Group Theory*, American Mathematical Society Translations, Ser 2, Vol. 9. American Mathematical Society, Providence, RI (1958), pp. 1–122.

[RS09]    K. RAFI and S. SCHLEIMER. Covers and the curve complex. *Geometric Topology*, (4)13 (2009), 2141–2162.

[Thu88]   W.P. THURSTON. On the geometry and dynamics of diffeomorphisms of surfaces. *Bulletin of the American Mathematical Society* (N.S.), (2)19 (1988), 417–431.

JING TAO, Department of Mathematics, University of Oklahoma, 601 Elm Avenue, Room 423, Norman, OK 73019-3103, USA                    jing@math.ou.edu