STRONGLY CONTRACTING GEODESICS IN OUTER SPACE

YAEL ALGOM-KFIR

Abstract. We study the Lipschitz metric on Outer Space and prove that fully irreducible elements of $\text{Out}(F_n)$ act by hyperbolic isometries with axes which are strongly contracting. As a corollary, we prove that the axes of fully irreducible automorphisms in the Cayley graph of $\text{Out}(F_n)$ are stable, meaning that a quasi-geodesic with endpoints on the axis stays within a bounded distance from the axis.

Introduction

There exists a striking analogy between the mapping class groups of surfaces, and the outer automorphism group $\text{Out}(F_n)$ of a rank n free group. At the core of this analogy lies Culler and Vogtmann's Outer Space CV_n [14], a contractible finite dimensional cell complex on which $\text{Out}(F_n)$ has a properly discontinuous action. Like Teichmüller space, Outer Space has an invariant spine on which the action is cocompact, making it a good topological model for the study of $\text{Out}(F_n)$. Indeed, Outer Space has played a key role in proving theorems for $\text{Out}(F_n)$, which were classically known for the mapping class group. For example, the action of a fully irreducible outer automorphism on the boundary of CV_n has been shown [21] to have North-South dynamics, and the Tits alternative holds for $\text{Out}(F_n)$ [5], [6].

However, while there have been several well studied metrics on Teichmüller space (the Teichmüller metric, the Weil-Petersson metric, and the Lipschitz metric), the geometry of Outer Space has remained largely uninvestigated (exceptions include [18] and [17]). One would like to define a metric on Outer Space so that fully irreducible elements of $\text{Out}(F_n)$, which are analogous to pseudo-Anosov elements in $\mathcal{MCG}(S)$, act by hyperbolic isometries with meaningful translation lengths. But immediately one encounters a problem: it isn't clear whether to require the metric to be symmetric. To clarify, we follow the discussion in [19]. Consider the situation of a pseudo-Anosov map ψ acting on Teichmüller space $\mathcal{T}(S)$ with the Teichmüller metric d_{T}. Associated to ψ is an expansion factor λ_ψ and two foliations $\mathcal{F}^s$ and $\mathcal{F}^u$ so that ψ expands the leaves of $\mathcal{F}^s$ by λ_ψ and contracts the leaves of $\mathcal{F}^u$ by λ_ψ^{-1}. Incidentally, $\lambda_\psi = \lambda_{\psi^{-1}}$. Furthermore, by Teichmüller 's theorem, the translation length of ψ is $\log(\lambda_\psi)$. Going back to $\text{Out}(F_n)$, one can associate to a fully irreducible outer automorphism Φ a Perron-Frobenius (PF) eigenvalue λ_Φ which plays much the same roll as the expansion factor in the pseudo-Anosov case. However, it is not always the case that $\lambda_\Phi = \lambda_{\Phi^{-1}}$. If we did have an honest metric on Outer Space where Φ was a hyperbolic isometry then the axis for Φ would also be an axis for Φ^{-1}. Thus for a point x on the axis of Φ, $d(x, \Phi(x)) = \log(\lambda_\Phi)$ and $d(\Phi(x), x) = \log(\lambda_{\Phi^{-1}})$. Therefore one would have to abandon either the symmetry of the metric or the relationship between the translation length of a fully irreducible element and its PF eigenvalue. We choose the former in order to preserve the ties between the action of Φ on CV_n and its action on the set of conjugacy classes in F_n.

The (non-symmetric) metric that we carry over from $\mathcal{T}(S)$ to Outer Space is the Lipschitz metric introduced by Thurston [27]. A map $h : X \to Y$ between two metric spaces is Lipschitz if there exists a constant L so that $d_Y(h(x), h(x')) \leq L d_X(x, x')$ for all $x, x' \in X$. The smallest L for which this holds is called the Lipschitz constant of h and denoted $\text{Lip}(h)$. Given two marked hyperbolic structures $(X, f), (Y, g)$ on a surface S define

$$d_L((X, f), (Y, g)) = \inf\{\text{Lip}(h) | h \text{ is Lipschitz, and homotopic to } g \circ f^{-1}\}$$

In [10] Choi and Rafi proved that this metric is Lipschitz equivalent to d_T in the thick part of $\mathcal{T}(S)$.

While $\mathcal{T}(S)$ with d_{T} is not CAT(0) [23] or Gromov hyperbolic [24] it does exhibit some features of negative curvature in the thick part. A geodesic is *strongly contracting* if its nearest point projection takes balls disjoint from the geodesic to sets of bounded diameter, where the bound is independent of the radius of the ball. Informally, the “shadow” that a ball casts on the geodesic is bounded. For example, geodesics in a Gromov hyperbolic space are strongly contracting. In [25], Minsky proved that geodesics contained in the ϵ-thick part of $\mathcal{T}(S)$ are uniformly strongly contracting, with the bound only depending on ϵ and the topology of S. Note that any axis of a pseudo-Anosov map is contained in the ϵ-thick part of $\mathcal{T}(S)$ for a sufficiently small ϵ. We prove

Theorem. *An axis of a fully irreducible outer automorphism is strongly contracting.*

A geodesic L in a metric space is *stable* if every quasi-geodesic segment with endpoints on L stays within a bounded neighborhood of L which only depends on the quasi-geodesic constants. As an application of the theorem above we prove:

Corollary. *In the Cayley graph of Out(F_n), the axis of a fully irreducible automorphism is stable.*

This paper is organized as follows

- In Chapter 1 we go over some definitions and background on Outer Space. The well informed reader could skip this part.
- In Chapter 2 we define the Lipschitz metric on Outer Space, and deduce a formula which expresses the relationship between the metric and the lengths of loops in X and Y (proof due to Tad White and first written in [17]). We show that the metric is symmetric up to a multiplicative constant in the ϵ-thick part of Outer Space.
- In Chapter 3 we describe axes of fully irreducible automorphisms. Given such an axis, we define a coarse projection of CV_n onto this axis. It is noteworthy that the axis for Φ will not necessarily be an axis for Φ^{-1}, however the distance of the projection of a point to $\mathcal{L}_\Phi$ to its projection to $\mathcal{L}_{\Phi^{-1}}$ does not depend on the point.
- In Chapter 4 we define the Whitehead graph $Wh_X(\Lambda^{\pm})$ of the attracting and repelling laminations of Φ at the point $X \in CV_n$. We prove that there

exists a point $F \in CV_n$ for which $Wh_F(\Lambda^+) \cup Wh_F(\Lambda^-)$ is connected and does not contain a cut vertex.

- In Chapter 5 we use the previous result to show that any loop α which represents a primitive conjugacy class cannot contain long pieces of both laminations. Next we prove our main "negative curvature" property. If the projections of x and y are sufficiently far apart then then $d(x, y)$ is coarsely larger than $d(x, p(x)) + d(p(x), p(y))$. We end the chapter by showing that this is enough to prove that $\mathcal{L}$ is a strongly contracting geodesic.
- In Chapter 6 we turn our attention to the Cayley graph of $\mathrm{Out}(F_n)$, proving that the axis of a fully irreducible automorphism is stable.
- In Chapter 7 we have collected some applications: the asymptotic cone of CV_n contains many cut points and is in fact tree graded, the divergence function in CV_n is at least quadratic, and we show that projections onto two axes A, B of independent irreducible automorphisms satisfy a dichotomy similar to the one shown in [2] for subsurface projections.

A note on notation: Many of the theorems and propositions in this article contain several constants which we usually denote s or c within the proposition. When referring to a constant from a previous proposition, we add its number as a subscript and state the defining property of this constant. We also remark that some of the figures contain color and are best viewed as such.

Acknowledgments. It is a pleasure to thank my advisor Mladen Bestvina for investing many hours of his time and a few of his brilliant ideas in this work. I would also like to thank Mark Feighn for his support and sound advice. Finally, I'd like to thank Kenneth Bromberg for motivating some of the applications in chapter 7.

1. Preliminary notions

Outer Space. Let F_n denote the free group on n letters. A rose R_0 is bouquet of n circles, with edges denoted by $e_1, \ldots, e_n$. Identify $\pi_1(R_0, \text{vertex})$ with $F_n =< x_1, \ldots, x_n >$ by declaring x_i to be the homotopy class of the loop e_i.
A metric graph is a graph with a geodesic metric (specified by assigning positive lengths to its edges). A marked graph (G, τ) is a graph G equipped with a homotopy equivalence $\tau : R_0 \to G$ called the *marking* of (G, τ) . τ induces an isomorphism $\tau_* : F_n \to \pi_1(G, \tau(\text{vertex}))$ identifying the fundamental group of G with F_n.
Culler and Vogtmann [14] defined Outer Space CV_n, as the space of equivalence classes of marked metric graphs (G, τ) where:

- each vertex in G has valence at least 3 and
- the equivalence relation is given by: $(G, \tau) \sim (H, \mu)$ if there exists a homothety $\rho : G \to H$ which preserves the marking up to homotopy, i.e. μ and $\rho \circ \tau$ are homotopic.

Another useful description of CV_n (which naturally extends to the boundary) is given in terms of F_n-trees. An $\mathbb{R}$-tree is a geodesic metric space in which every two points x, y are connected by a unique embedded path. A point $p \in T$ is a branch point if $T \setminus p$ has three or more connected components. An $\mathbb{R}$-tree with an isometric F_n action is called an F_n-tree. An F_n-tree is simplicial if the set of branch points is discrete, it is minimal if it has no invariant subtree. Outer Space is the space of equivalence classes of free, simplicial, minimal F_n-trees where two trees T, T' are equivalent if there exists an equivariant homothety $\rho : T \to T'$. The functor taking each graph G to its universal cover $\widetilde{G}$ induces an equivalence of the two definitions of Outer Space.

The axes topology. Consider the set of non-trivial conjugacy classes $\mathcal{C}$ in F_n. Each F_n-tree T induces a length function $\ell_T : F_n \to \mathbb{R}$ by $\ell_T(x) =$ the translation length $\mathrm{tr}(x)$ of x as an isometry of T. Since the translation length is a class function, ℓ_T descends to a map $\ell_T : \mathcal{C} \to \mathbb{R}$. Therefore we can define a map

$$\begin{array}{rcl} \ell : CV_n & \to & \mathbb{RP}^{\mathcal{C}} \\ {[T]} & \to & [\ell_T] \end{array}$$

In [13] Culler and Morgan proved that this map is injective. Thus CV_n inherits a topology from $\mathbb{RP}^{\mathcal{C}}$ known as the axes topology. We remark (although we will not need this) that there are other ways to define a topology on CV_n: using the cellular structure of CV_n, and using the Gromov topology on the space of metric F_n-trees. Paulin [26] proved that all three topologies are equivalent.

The boundary of Outer Space. In [13] Culler and Morgan showed that $\overline{CV_n}$ is compact. It was later shown in [11] and [3] that $\overline{CV_n}$ is the space of very small minimal F_n-trees up to equivalence by equivariant homothety.

Free factors, basis elements and Whitehead's theorem. A free factor is a subgroup $A < F_n$ for which there is a $B < F_n$ such that $F_n = A * B$. An element $x \in F_n$ is called a *basis element* (or a primitive element) if $< x >$ is a free factor. For example, no commutator is a basis element of F_n because it represents the trivial element in the abelianization of F_n. In [28], Whitehead discusses an algorithm for deciding if a set $[y_1], \dots, [y_k]$ of conjugacy classes can be completed to a basis, i.e. do there exist elements $w_1, \dots, w_k$ such that $y_1^{w_1} \dots y_k^{w_k}$ can be completed to a basis of F_n. We will need a version of his theorem here.

Definition 1.1. Let $[y]$ be the conjugacy class of a cyclically reduced word y in F_n written in the generators $x_1, \dots, x_n$. Then the Whitehead graph of $[y]$ with respect to the basis $\mathcal{B} = \{x_1, \dots, x_n\}$ is denoted $Wh_{\mathcal{B}}([y])$ and constructed as follows: The vertex set of this graph is the set $\mathcal{B} \cup \mathcal{B}^{-1}$. $z_i, z_j \in \mathcal{B} \cup \mathcal{B}^{-1}$ are connected by an edge if $z_i^{-1} z_j$ appears in the cyclic word y. Equivalently, let Y is the 2-complex constructed by gluing a disc D to R_0 via the attaching map $\partial D \xrightarrow{[y]} R_0$ then $Wh_{\mathcal{B}}([y]) = \mathrm{Link}(\mathrm{ver}, Y)$. The Whitehead graph of the set of conjugacy classes $[y_1], \dots, [y_k]$ is the union of all the individual whitehead graphs, taken with the same vertex set $Wh_{\mathcal{B}}([y_1], \dots, [y_k]) = \cup_{i=1}^k Wh_{\mathcal{B}}([y_i])$.

Whitehead proves

Theorem 1.2 ([28])**.** *If $[y_1], \dots, [y_k]$ can be completed to a basis, and for some basis $\mathcal{B}$, $Wh_{\mathcal{B}}([y_1], \dots, [y_k])$ is connected. Then $Wh_{\mathcal{B}}([y_1], \dots, [y_k])$ contains a cut vertex.*

If there's a cut vertex a in $W = Wh_{\mathcal{B}}([y_1], \dots, [y_k])$ then one could decrease $\sum_{i=1}^k |y_i|_{\mathcal{B}}$ by changing the basis as follows. Consider W' the subgraph of W induced by the set of vertices different from a. Let W^0 be the vertices in the connected component of the vertex a^{-1} in W', and let W^1 be the rest of the vertices in W'. Construct the basis $\mathcal{B}_1$ from $\mathcal{B}$:

- if both $x_j, x_j^{-1} \in W^1$ then replace x_j with $x_i x_j x_i^{-1}$,
- if only $x_j \in W^1$ then replace x_j with $x_i x_j$
- if both $x_j, x_j^{-1} \in W^0$ then leave them as they are.

It is straightforward to check that $\mathcal{B}_1$ is indeed a basis for F_n and that $\sum_{i=1}^k |y_i|_{\mathcal{B}} > \sum_{i=1}^k |y_i|_{\mathcal{B}_1}$. One could repeat the process to obtain a fast algorithm for deciding if $[y_1], \dots, [y_k]$ are part of a basis.

We will need a slightly different version of Whiteheads theorem, which can be found in R. Martin's PhD thesis [22] (proof attributed to M. Bestvina). The theorem in [22] is stated and proved for one conjugacy class but the proof of the statement for several conjugacy classes needs no modification thus we omit it. To state the theorem we need to define the generalized Whitehead graph and the decomposition space of $[y_1], \dots, [y_k]$.

We define the generalized Whitehead graph $Wh_m([y_1], \dots, [y_k]; \mathcal{B})$ as follows. Consider R_0 the bouquet of n circles $e_1, \dots, e_n$ labelled by $x_1, \dots, x_n$. Represent $[y_1], \dots, [y_k]$ by immersed loops $\alpha_1, \dots, \alpha_k$ in R_0. Let $T = \widetilde{R_0}$ the universal cover of R_0 with some basepoint, and lift $\alpha_1, \dots, \alpha_k$ to a family of lines in T. The vertices of the Whitehead graph are in 1-1 correspondence with the vertices lying on the sphere of radius m centered at the basepoint, i.e. the words of length m in the basis $x_1, \dots x_n$. Two of these vertices w, w' are connected by an edge in $Wh_m([y_1], \dots, [y_k]; \mathcal{B})$ if there exists $\widetilde{\alpha_i}$ a lift of one of the α_is such that w, w' both lie on $\widetilde{\alpha_i}$. It is easy to see that $Wh_1([y_1], \dots, [y_k]; \mathcal{B}) = Wh_{\mathcal{B}}([y_1], \dots, [y_k])$.

We will also associate to $[y_1], \dots [y_k]$ a quotient space of ∂T. The boundary of T is the set of ends of T, i.e. the set of equivalence classes of rays where two rays are identified if they have infinite intersection. A basis for a topology on ∂T is given by finite words in F_n: Let w be a reduced word in F_n and let β be the path from 1 to w . U_w is the set of geodesic rays emanating from the basepoint that contain β. $\{U_w\}$ is a basis for a topology on ∂T. Given $[y_1], \dots, [y_k]$ define $\partial T/[y_1], \dots, [y_k]$ to be the quotient space of ∂T where $[r] \sim [r']$ if there is an $\widetilde{\alpha}$ a lift of one of $\alpha_1, \dots, \alpha_k$ such that one of $\widetilde{\alpha}$ endpoints is $[r]$ and the other is $[r']$. Endow $\partial T/\sim$ with the quotient topology. $\partial T/[y_1], \dots, [y_k]$ is called the decomposition space associated to $[y_1], \dots, [y_k]$.
We are now ready to reformulate Whitehead's result:

Theorem 1.3 ([22]). *For any set of elements $y_1, y_2, \dots, y_k$ the following are equivalent:*

1. *$[y_1], \dots, [y_k]$ agree with a proper splitting of F_n up to conjugacy, i.e. there is a splitting $F_n = A * B$ such that for each $1 \le i \le k$: $[y_i] \in [A]$ or $[y_i] \in [B]$.*
2. *The decomposition space $\partial F_n/[y_1], \dots, [y_k]$ is disconnected.*
3. *For any basis $\mathcal{B}$ there is some m such that $Wh_m([y_1], \dots, [y_k]; \mathcal{B})$ is disconnected.*
4. *If $\mathcal{B}$ is a basis such that $Wh_{\mathcal{B}}([y_1], \dots, [y_k])$ contains no cut vertex then it is disconnected.*
5. *There exists a basis $\mathcal{B}$ such that $Wh_{\mathcal{B}}([y_1], \dots, [y_k])$ is disconnected.*

Outer automorphisms. While it has now become an object of independent study, initially Outer Space's raison d'être was the study of the group of outer automorphisms of the free group $\mathrm{Out}(F_n) = \mathrm{Aut}(F_n)/\mathrm{Inn}(F_n)$. An outer automorphism Φ is reducible if there is a splitting $F_n = A_1 * A_2 * \cdots * A_k * B$ with $A_i \neq 1$ and either $B \neq 1$ or $k \geq 2$ such that Φ permutes the factors $A_1, \ldots, A_k$ up to conjugacy. If Φ is not reducible then it is *irreducible*. Φ is *fully irreducible*, also called IWIP (irreducible with irreducible powers), if for all $i > 0$, Φ^i is irreducible. Φ is *geometric* if there exists a surface automorphism $f : S \to S$ of a surface S with $\pi_1(S) = F_n$ such that $[f_*] = \Phi$. Except for the case $n = 2$ where all outer automorphisms are geometric; geometric outer automorphisms are a small subset of $\mathrm{Out}(F_n)$.
There is a natural right action of $\mathrm{Out}F_n$ by homeomorphisms of CV_n (when we introduce a metric on CV_n we shall see that this is actually an isometric action). Let $(G, h) \in CV_n$ and Φ an outer automorphism. Let $f : R_0 \to R_0$ be a map such that $[f_*] = \phi$ define $(G, h) \cdot \Phi = (G, h \circ f)$. The homotopy class of $h \circ f$ is independent of the choice of f so we get a well defined action. This action has finite stabilizers but is not cocompact. However, CV_n has a *spine* (a contractible deformation retract of CV_n) on which the restriction of the action of $\mathrm{Out}(F_n)$ is cocompact (see [14] for details). Therefore, CV_n is a nice topological model for $\mathrm{Out}(F_n)$. We aim to demonstrate in this work that it is a good geometric model for it as well.

Train-track structures and maps. Let G be a graph. An unordered pair of oriented edges $\{e_1, e_2\}$ is a *turn* if e_1, e_2 have the same initial endpoint. Let $\bar{e}$ denote the edge e with the opposite orientation. If an edge path $\alpha = \cdots \overline{e_1} e_2 \cdots$ or $\alpha = \cdots \overline{e_2} e_1 \cdots$ then we say that α *crosses* or *contains* the turn $\{e_1, e_2\}$.

Definition 1.4. A *train track structure* on G is an equivalence relation on the set of oriented edges $E(G)$ with the property that if $e_1 \sim e_2$ then $\{e_1, e_2\}$ is a turn.

A turn $\{e_1, e_2\}$ is *legal* with respect to a fixed train-track structure on G if $e_1 \nsim e_2$. An edge path is legal if every turn it crosses is legal. The equivalence classes of the edges are called *gates*. Let $f : G \to G$ be a map such that on every edge it either collapses it to a point, or restricts to an immersion. Let $Df : E \to E$ be a map on E - the set of oriented edges in G, which sends each edge to the first edge in its image (can be thought of as a kind of derivative). f induces a map Tf on the set of turns $Tf(e, e') = (Df(e), Df(e'))$. In [8] Bestvina and Handel made the following definition.

Definition 1.5. f is called a *train-track map* if it preserves a train-track structure, i.e. if there is a train-track structure on G such that f takes legal turns to legal turns.
Equivalently, f is a train-track map if $f^i(e)$ is an immersed path for each $i \in \mathbb{N}$ and each edge $e \in E$.

In [8] they lay out an algorithm which produces train-track map f for any irreducible outer automorphism Φ, with $[f_*] = \phi$. Up to scale, there is a unique assignment of lengths to G and a constant λ (the growth rate of f, or the PF eigenvalue of f) so that for every edge e we have $length(f(e)) = \lambda length(e)$. Although f is not unique in the sense that there might be many train-track maps representing a given automorphism, λ depends only on Φ.
Given a concatenation of legal paths $\alpha \cdot \beta$, there might be some cancellation in $f(\alpha) \cdot f(\beta)$. However, Cooper proved in [12] that the amount of cancellation bounded

by a constant $K = BCC(f)$ which depends only on f (and not on the paths). Following [4] we define C_{crit} the *critical constant* for f as $C_{\text{crit}} = \frac{2K}{\lambda - 1}$ (where λ is the PF eigenvalue of f). For every $C > C_{\text{crit}}$ there is a μ such that for any legal paths α, β, γ where β's length is greater than C, $[f^k(\alpha \cdot \beta \cdot \gamma)]$ contains a legal segment of length at least $\mu\lambda^k \text{length}(\beta)$ contributed by $f^k(\beta)$.

Laminations of fully irreducible automorphisms. Let $f : G \to G$ be a train-track map of an irreducible outer automorphism. By replacing f with a power if necessary, we may assume that f has a fixed point x in the interior of an edge. Let I be an ϵ neighborhood of x so that $f(I) \supset I$. Choose an isometry $\ell : (-\epsilon, \epsilon) \to I$ and extend uniquely to a local isometric immersion $\ell : \mathbb{R} \to G$ so that $\ell(\lambda^m t) = f^m(t)$ for all $t \in \mathbb{R}$. The attracting lamination in G, $\Lambda^+(G)$, is the collection of all such immersions ℓ called attracting leaves. A stable leaf subsegment is the restriction of ℓ to a subinterval of $\mathbb{R}$. Different leaves have the same leaf subsegments. Given a different metric graph $H \in CV_n$ and a homotopy equivalence $\tau : G \to H$, $\Lambda^+(H)$ the attracting lamination in the H coordinates is the collection of immersions $[\tau\ell]$ pulled tight. An important feature of these laminations is:

Proposition 1.6 ([4] Proposition 1.8)**.** *Every leaf of Λ^+_Φ is quasi-periodic.*

This means that for every length L there is a length L' such that if $\alpha, \beta \subseteq \ell$ are subleaf segments with $\mathit{length}(\alpha) = L$ and $\mathit{length}(\beta) > L'$ then β contains an occurrence of α. One can think of ℓ as a necklace made of beads. The segments of length L that appear in ℓ are beads of different colors. The proposition tells us that in any subchain of $\frac{L'}{L}$ consecutive beads we can find beads of all possible colors.

Parameterization of Outer Space. Unnormalized Outer Space $\mathcal{X}$ is the space of metric graphs where two are equivalent if there is an isometry homotopic to the difference in marking. There is a similar definition in terms of trees. CV_n is a section of $\mathcal{X}$. One parameterization of CV_n is to take the set of volume 1 graphs in $\mathcal{X}$

This is useful, but unfortunately, doesn't extend to the boundary. Another parameterization that we will use which does extend to the boundary is as follows: By Serre's theorem there is a finite set of conjugacy classes (for example $[x_1], \dots, [x_n], [x_1x_2], \dots [x_{n-1}x_n]$) which cannot be simultaneously elliptic in any minimal F_n-tree. Then CV_n is the section where the sum of their translation lengths is 1.

2. The Lipschitz metric on CV_n

Let $(X, f), (Y, g)$ be two marked metric graphs of volume 1. Consider:

$$L(X, Y) = \left\{ \text{Lip}(h) \in \mathbb{R} \,\middle|\, h \text{ is Lipschitz and homotopic to } g \circ f^{-1} \right\}$$

Note that h does not necessarily take vertices to vertices. Define the distance from X to Y to be:

$$d(X, Y) = \log \inf L(X, Y)$$

By Arzela-Ascoli $L(X, Y)$ has a minimum. Moreover, it is enough to consider maps which are linear on edges. Indeed, for any map h one can construct a homotopic map h_1 which is linear on edges by defining $h_1(v) = h(v)$ on every vertex v and sending an edge (v, w) to the immersed path $[h(v), h(w)]$ which is homotopic to

$\text{Im}(h|_{(v,w)})$ rel endpoints and parameterized at constant speed. It is clear that $\text{Lip}(h_1) \leq \text{Lip}(h)$. Therefore we can usually restrict our attention to such maps.

To see that $d(X,Y)$ really defines a (non-symmetric) metric note the following:

(1) If $d(X,Y) = 0$ then there is a map h homotopic to the difference in marking with $\text{Lip}(h) = 1$. Since h is a homotopy equivalence and X, Y don't contain valence 1 vertices then h is surjective. $\text{Lip}(h) = 1$ implies that h doesn't stretch edges. Since h is onto and $vol(X) = vol(Y) = 1$, it cannot shrink edges either. h is linear so $h'(x) = 1$ for all $x \in X$. h is injective, otherwise $vol(Y) \leq vol(\text{Im}h) < vol(X)$. We conclude that h is a bijection with $h' \equiv 1$ thus it is an isometry from X to Y. Since it is homotopic to the difference in marking we get $X = Y$ in CV_n.

(2) If $h_1 : X \to Y$, $h_2 : Y \to Z$ are optimal maps, then $h = h_2 \circ h_1$ is homotopic to the difference in marking from X to Z, thus $\text{Lip}(h) \geq \min L(X,Z)$. But the chain rule implies $\text{Lip}(h_1)\text{Lip}(h_2) \geq \text{Lip}(h) \geq \min L(X,Z)$ taking log we get $d(X,Y) + d(Y,Z) \geq d(X,Z)$.

Conventions and notation. If α_X is an immersed loop in X then it represents a conjugacy class α of $\pi_1(X,p)$. For every marked graph Z, we shall denote by α_Z the *immersed loop* representing α in the graph Z.
If the loop α_X is not an immersed loop, we denote by $[\alpha_X]$ the immersed loop representing the same conjugacy class.
For a path γ_X in $X \in CV_n$, the length of γ_X in X is denoted by $l(\gamma_X, X)$. For a conjugacy class γ in F_n the length of the immersed loop representing γ in X will be denoted $l(\gamma, X)$.

We now prove that the optimal Lipschitz constant of a map from X to Y is equal to the stretch factor of the maximally stretched loop. For two points $X, Y \in CV_n$ and a conjugacy class α, denote the stretch of α from X to Y by $\text{St}_\alpha(X,Y) = \frac{l(\alpha,Y)}{l(\alpha,X)}$. Let

$$S(X,Y) = \{\text{St}_\alpha(X,Y) |\, \alpha \text{ is a conjugacy class in } F_n\}$$

Theorem 2.1 (T. White see [17])**.**

$$\min L(X,Y) = \max S(X,Y)$$

Proof. Let α be a conjugacy class represented in X by the immersed loop α_X. The loop $h(\alpha_X)$ represents α in Y, it might not be immersed. We denote by $[h(\alpha_X)]$ the tightened image of $h(\alpha_X)$. Then

$$l(\alpha, Y) = l([h(\alpha_X)], Y) \leq l(h(\alpha_X), Y) \leq \text{Lip}(h) l(\alpha, X) \tag{1}$$

We deduce that $\max S(X,Y) \leq \min L(X,Y)$. Notice that we get equality in equation 1 iff all of the edges which α_X crosses are stretched by $\text{Lip}(h)$ and $h(\alpha_X)$ is a tight loop in Y. The goal is now to find a map h for which such a loop exists.

Given a map $h : X \to Y$ which is linear on edges, let X_h be the subgraph of all edges e such that $\text{St}(h,e) = \text{Lip}(h)$. h induces a train-track structure on X: two edges e_1, e_2 belong to the same gate if $h(e_1), h(e_2)$ define the same germ. The inclusion $X_h \subseteq X$ induces a train-track structure on X_h.

Let h be an optimal map, linear on edges and so that X_h is smallest among all optimal maps. We claim there are at least two gates at each vertex. See figure 1 for an example. In general, suppose by way of contradiction that there is a vertex v where X_h contains only one gate at v. Let $S_2 = \max\{\text{St}(h,e) | e \notin X_h\}$ be the second largest derivative, and $\varepsilon = \mathrm{Lip}(h) - S_2$. Define a map h_1 by $h_1(u) = h(u)$ for all vertices $u \neq v$. To define $h_1(v)$ take any $e \in X_h$ adjacent to v, define $h_1(v)$ the point on the germ defined by $h(e)$ a distance $< \varepsilon \cdot$ (length of smallest edge in Y) away from $h(v)$. Define h_1 to be homotopic to h and linear on edges.

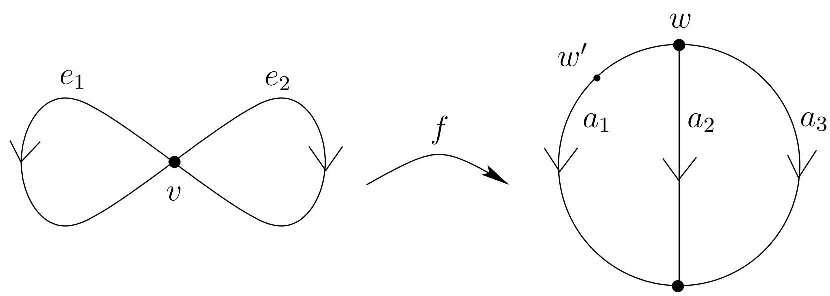

FIGURE 1. The following is an example of a map f where X_f has a vertex with one gate. We show that the map is not optimal. In the graph on the left both edges have length $\frac{1}{2}$ and in the graph on the right all three edges have length $\frac{1}{3}$. Suppose the map f takes $e_1 \to a_2\overline{a_3}$ and $e_2 \to a_1\overline{a_2}a_1\overline{a_3}a_2\overline{a_1}$. The stretch of f on e_1 is $\frac{2/3}{1/2} = \frac{4}{3}$ and the stretch of f on e_2 is $\frac{6/3}{1/2} = 4$, so $\mathrm{Lip}(f) = 4$ and $X_f = \{e_2, \overline{e_2}\}$. Both $e_2, \overline{e_2}$ begin with a_1 so X_f contains only one gate at v. Let w' be a point on a_1 which is ε away from v where $2\varepsilon < 4 - \frac{4}{3}$. w' divides a_1 into two edges b_1, b_2. Consider the map f_1 which takes $e_1 \to \overline{b_1}a_2\overline{a_3}b_1$ and $e_2 \to b_2\overline{a_2}b_1b_2\overline{a_3}a_2\overline{b_2}$. f_1 is homotopic to f. f_1 stretches e_1 by $\frac{2/3+2\varepsilon}{1/2} = \frac{4+12\varepsilon}{3}$ and e_2 by $\frac{2-2\varepsilon}{1/2} = 4 - 4\varepsilon$. Since ε is small enough $\mathrm{Lip}(f_1) = 4 - 4\varepsilon$ which is smaller than $\mathrm{Lip}(f)$.

We show that $S_2 < \mathrm{St}(h_1, e) < \mathrm{Lip}(h)$ for every edge $e \in X_h$ adjacent to v. $\mathrm{St}(h_1, e) = \frac{\mathrm{Lip}(h)len(e) - \varepsilon \cdot \text{length of smallest edge}}{len(e)} > \frac{\mathrm{Lip}(h)len(e) - \varepsilon len(e)}{len(e)} = \mathrm{Lip}(h) - \varepsilon = S_2$. Therefore, either $\mathrm{Lip}(h_1) = \mathrm{Lip}(h)$ with $X_{h_1} \subsetneq X_h$, or $\mathrm{Lip}(h_1) < \mathrm{Lip}(h)$ which contradictions the choice the h. Thus we've shown that there are at least two gates for every vertex of X_h.

Consider a legal path $\alpha_X \in X_h$ which intersects itself twice. Such a path will contain a legal subloop. Indeed parameterize α so that $\alpha : [0,1] \to X_h$ with $\alpha(0) = \alpha(t_1)$, and $\alpha(t_2) = \alpha(1)$ where $0 < t_1 \leq t_2 < 1$. If $D_+\alpha(0) \neq D_-\alpha(t_1)$ then $\alpha|_{[0,t_1]}$ is a legal loop, if $D_+\alpha(t_2) \neq D_-\alpha(1)$ then $\alpha|_{[t_2,1]}$ is a legal loop. Otherwise $\alpha * [\alpha(t_2), \alpha(t_1)]$ is a legal loop. Legal loops are mapped to tight loops in Y thus if h is as above there is a maximally stretched loop α with stretch constant equal to the Lipschitz constant of h. □

This proof shows a bit more: there is a relatively short loop which is maximally stretched. The following theorem from [17] makes this precise.

Theorem 2.2 ([17]). *For any two points $X, Y \in CV_n$ there is a maximally stretched loop α_X whose image is one of the following:*

(1) *an embedded circle.*
(2) *a wedge of two embedded circles.*
(3) *a barbell i.e.* α *is a concatenation of* $\gamma_1\gamma_2\gamma_3\overline{\gamma_2}$ *where* γ_1 *and* γ_3 *are disjoint embedded circles and* γ_2 *is an embedded path which intersects* γ_i, $i = 1, 3$ *in exactly one point which is one of its endpoints.*

Such a loop can contain at most $|E|$ edges where $|E|$ is the number of edges in X. By an Euler-characteristic argument $|E| \leq 3n-3$ therefore α_X crosses at most $3n-3$ edges.

Definition 2.3. Let X be a point in CV_n. Define the set of candidates of X, $\mathrm{Can}(X)$ as the set of loops of the type described in theorem 2.2.

Corollary 2.4. *The distance is realized by one of the candidates of* X*:*

$$d(X,Y) = \log\max\left\{ \frac{l(\alpha,Y)}{l(\alpha,X)} \middle| \alpha \in \mathrm{Can}(X) \right\}$$

Corollary 2.5. *If* $\alpha_X, \beta_X \in \mathrm{Can}(X)$ *and their images in* X *are different sets then* α, β *can be completed to a basis.*

Proof. Let J by a maximal forest in X which doesn't contain an edge e_i of $\mathrm{Im}\alpha_X \setminus \mathrm{Im}\alpha_Y$. Collapse J to get R_X a wedge of circles. Since e_i was not collapsed, $\mathrm{Im}\alpha_R \neq \mathrm{Im}\beta_R$. Let $e_j \subseteq \beta_R$ be any edge then $< \alpha_R, \beta_R, e_1, \ldots, \hat{e_i}, \ldots, \hat{e_j}, \ldots e_n >$ represents a basis for F_n. □

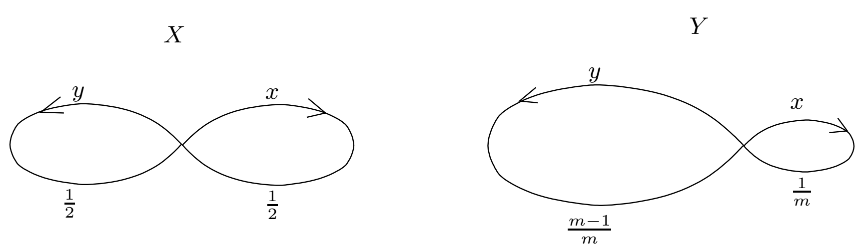

FIGURE 2. An example where $d(X,Y) = \log\frac{2m-2}{m} \sim \log 2$ and $d(Y,X) = \log\frac{m}{2}$

2.1. Non-symmetry of $d(X,Y)$.

Example 2.6. Figure 2 shows an example of two graphs X, Y_m such that $d(X,Y) = \log \mathrm{St}_y(X,Y) = \log\left(\frac{2m-2}{m}\right)$ however going the other way, the distance is attained on x and $d(Y,X) = \log \mathrm{St}_x(X,Y) = \log\left(\frac{1/2}{1/m}\right) = \log\left(\frac{m}{2}\right)$. Therefore we cannot hope for the metric to be symmetric up to a multiplicative constant.

However, not all is lost, one can say something if we forget the metric structure on the graphs and consider them as topological graphs with a marking. In [18] Handel and Mosher studied a combinatorial metric on the space of such graphs defined as follows. Let $h: X \to Y$ be a simplicial map (i.e. vertices are mapped to vertices) homotopic to the difference in marking. Define the stretch of h as the total volume of the image of h, i.e. $T(h) = \sum_{e\in X} |h(e)|$ where $|h(e)|$ is the number of edges $h(e)$ crosses. Let $d_{\mathrm{comb}}(X,Y) = \log\min\{T(h) | h \sim g \circ f^{-1}\}$. They prove:

Theorem 2.7 ([19])**.** *There is a constant $c = c(n)$ such that*

$$\frac{1}{c} \cdot d_{\text{comb}}(Y, X) \le d_{\text{comb}}(X, Y) \le c \cdot d_{\text{comb}}(Y, X)$$

for any $X, Y \in CV_n$.

Definition 2.8. For $\theta > 0$ the θ-thick part of CV_n is

$$CV_n(\theta) = \{X \in CV_n | l(\alpha, X) \ge \theta \text{ for all } \alpha\}$$

We can relate the Lipschitz metric to the combinatorial metric in $CV_n(\theta)$.

Lemma 2.9. *For any $\theta > 0$ there is constant $c = c(\theta)$ such that: $d_{\text{comb}}(X, Y) - c \le d(X, Y) \le d_{\text{comb}}(X, Y) + c$ for all $X, Y \in CV_n(\theta)$.*

Proof. Notice that each graph has no more than $3n - 3$ edges, and the length of any edge is bounded below by θ and above by $1 - \theta$. Let $T(h) = e^{d_{\text{comb}}(X,Y)}$ then

$$\text{Lip}(h) = \max\left\{ \frac{l(h(e),Y)}{l(e,X)} \middle| e \in X \right\} \le \sum_{e\in X} \frac{l(h(e),Y)}{l(e,X)} \le \sum_{e\in X} \frac{(1-\theta)|h(e)|}{\theta}$$
$$\frac{(1-\theta)}{\theta} \sum_{e\in X} |h(e)| = \frac{1-\theta}{\theta} T(h)$$

Therefore $d(X, Y) \le d_{\text{comb}}(X, Y) + \log \frac{1-\theta}{\theta}$.
Now suppose that $f : X \to Y$ realizes the Lipschitz distance. Homotope f to a simplicial map $f' : X \to Y$ by moving the images of vertices in X to vertices of Y, and pulling tight the images of edges. For each edge $l(f'(e), Y) \le l(f(e), Y) + 2(1 - \theta)$. Thus

$$e^{d(X,Y)} = \text{Lip}(f) = \max\left\{ \frac{l(f(e),Y)}{l(e,X)} \middle| e \in X \right\} \ge \max\{l(f(e), Y)) \mid e \in X\} \ge$$
$$\frac{1}{3n-3} \sum_{e\in X} l(f(e), Y) \ge \frac{1}{3n-3} \sum_{e\in X} [l(f'(e), Y) - 2(1-\theta)] \ge$$
$$\frac{\theta}{3n-3} \sum_{e\in X} |f'(e)| - 2(1-\theta) \ge \frac{\theta}{2(3n-3)} T(f')$$

The last inequality holds when $T(f') \ge \frac{4(3n-3)}{\theta}$. We thus get

$$d(X, Y) \ge d_{\text{comb}}(X, Y) - \log \frac{2(3n-3)}{\theta}$$

□

Putting together Lemma 2.9 and Theorem 2.7 we get:

Corollary 2.10. *For any $\theta > 0$ there is constant $c = c(\theta)$ such that:*

$$\frac{1}{c} \cdot d(Y, X) \le d(X, Y) \le c \cdot d(Y, X)$$

for any $X, Y \in CV_n(\theta)$.

3. Axes in CV_n of fully irreducible automorphisms

It is straightforward to check that the right action of $\text{Out}(F_n)$ on CV_n is an isometric action. Furthermore, if $f : G \to G$ represents $\phi \in \text{Out}(F_n)$ and the loop α_G represents α in G then $[f(\alpha_G)]$ is a tight loop in $G\phi$ representing α in $G\phi$.

Definition 3.1 (Fold line)**.** Given an irreducible train-track map $f : G_0 \to G_0$ we define a fold line from G_0 to $G_0 \cdot \phi$. At each stage we will have a map $g_s : G_s \to G_0 \cdot \phi$ where $g_0 = f$ and $g_{\log(\lambda)} = id$. We define the path $G : [0, \log(\lambda)] \to CV_n$ and the maps g_s inductively. Progress is measured by the size of $P(g_s) = g_s^{-1}(\text{vertices of } G_0\phi)$. Let e_1, e_2 be two edges, with common initial vertex v and which are taken to the same germ under f, i.e. $\{e_1, e_2\}$ is illegal. Consider e_i as paths with $e_i(0) = v$ and let $R = \max\{r \mid \forall s \leq r : f(e_1(s)) = f(e_2(s))\}$. At time t, G_t is the graph obtained from G_0 by folding the initial subsegments of e_1, e_2 of length $r = 1 - e^{-t}$. Formally, G_t is the quotient graph of G_0 under the equivalence relation $e_1(s) = e_2(s)$ for all $s \leq r$ where $r = 1 - e^{-t}$ and where edge lengths are changed as follows: the length of all edges other than e_1, e_2 are scaled by e^t, and the length of e_i in G_t is $e^t(|e_i|_0 - (1 - e^{-t}))$. We get a path $G : [0, t_1] \to CV_n$, where $t_1 = \log\left(\frac{1}{1-R}\right)$. We will later show that this is a geodesic parameterized according to arc length. We define $g_s : G_s \to G_0 \cdot \phi$ as the quotient map of $f : G_0 \to G_0 \cdot \phi$ under the equivalence relation. Notice that $|P(g_{t_1})| \leq |P(g_0)| - 1$. We continue constructing the path using g_{t_1} instead of f. We're guaranteed to stop after $|P(f)|$ steps, obtaining a fold path G from G_0 to $G_0 \cdot \phi$. Notice that we have made choices when deciding the order in which to fold the edges, but in any case this construction yields at least one fold line.

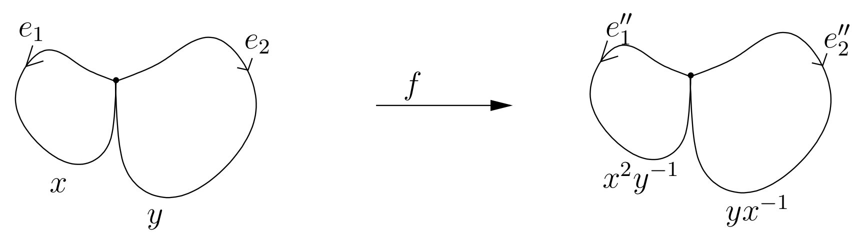

FIGURE 3. A train track graph G for the map f in example 3.2 and its image under the action of ϕ. The edge lengths in $G(0)$ are $l(e_1) = a = \frac{3-\sqrt{5}}{2} \sim 0.382$ and $l(e_2) = b = \frac{\sqrt{5}-1}{2} \sim 0.618$, $\lambda \sim 2.618$.

Example 3.2. Consider the automorphism

$$\phi = \begin{cases} x \longrightarrow xy \\ y \longrightarrow yxy \end{cases}$$

Since it is a positive automorphism the corresponding map on the rose with two petals e_1, e_2 is a train-track map. To find the lengths of e_1, e_2 and the PF eigenvalue λ we solve the linear system of equations

$$\begin{cases} a + b = \lambda a \\ a + 2b = \lambda b \\ a + b = 1 \end{cases}$$

We get that the length of e_1 is $a = \frac{3-\sqrt{5}}{2}$, the length of e_2 is $b = \frac{\sqrt{5}-1}{2}$ and $\lambda = \frac{3+\sqrt{5}}{2}$. See figure 3 for a picture of $G_0\phi$, the labels on the edges denote the inverse of the marking. Notice that for each loop α in G_0, $f(\alpha)$ represents the same conjugacy class in $G_0\phi$. Figure 4 shows a fold path from G_0 to $G_0\phi$. We start by folding $\overline{e_1}, \overline{e_2}$ until we completely wrap e_2 over e_1. We get a new graph $G(\log \frac{1}{1-a}) = G(\log \frac{\lambda}{\lambda-1})$, with edges e_1', e_2' and edge lengths $\frac{a}{b} = b$ and $\frac{b-a}{b} = a$ respectively. Now we fold $\overline{e_1'}$ onto $\overline{e_2'}$ until we completely wrap e_1' over e_2'. The parameter t of the final graph is $\log(\frac{\lambda}{\lambda-1}) + \log(\lambda - 1) = \log(\lambda)$ and we get $G(\log(\lambda)) = G(0)\phi$.

Proposition 3.3. *The fold line $G : [0, \log(\lambda)] \to CV_n$ is a geodesic parameterized according to arc length.*

Proof. Let $s \in [0, \log(\lambda)]$ and define $h_s : G_0 \to G_s$ to be the composition of quotient maps so that $f = g_s \circ h_s$. The key observation is that if α is legal in G_0 with respect to f then it is legal in G_0 with respect to h_s and $h_s(\alpha)$ is legal in G_s with respect to g_s (see figure 5). Indeed consider $\alpha : \mathbb{S}^1 \to G_0$. α is illegal with respect to h_s if there is a point $x \in \mathbb{S}^1$ and an ε-neighborhood of x such that α maps a punctured ε-neighborhood of x homeomorphically into G_0 and is two-to-one into G_s. However it must then be at least two-to-one in $G_{\log(\lambda)}$ so it is illegal with respect to f. Similarly, if $h(\alpha)$ is illegal with respect to g_s then α is illegal with respect to f.

Recall that $d(X, Y) = \log \mathrm{St}_\alpha(X, Y)$ for any legal loop α with respect to the train track structure induced by the difference in marking. Let β be a legal loop in G_0 with respect to f then it is legal with respect to h_s and $h_s(\beta)$ is legal with respect to g_s. Thus $d(G_0, G_0 \cdot \phi) = \log \mathrm{St}_\beta(G_0, G_0 \cdot \phi)$, $d(G_0, G_s) = \log \mathrm{St}_\beta(G_0, G_s)$, $d(G_s, G_0 \cdot \phi) = \log \mathrm{St}_{h(\beta)}(G_s, G_0 \cdot \phi)$. Moreover, since there is no backtracking:

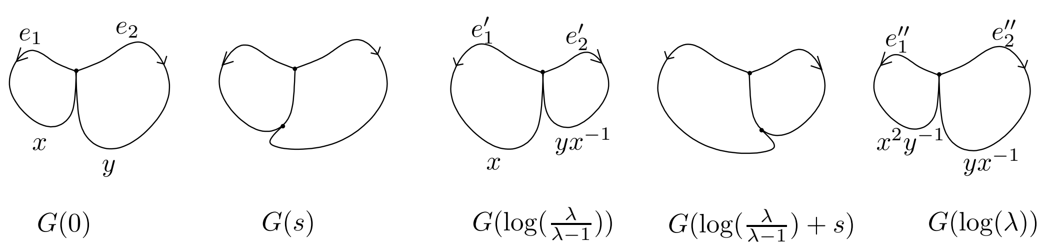

FIGURE 4. A fold path $G : [0, \log \lambda] \to CV_n$. Notice that if a loop is legal in $G(0)$ then it remains legal in $G(s)$. This can be used to show that the path is a geodesic in CV_n

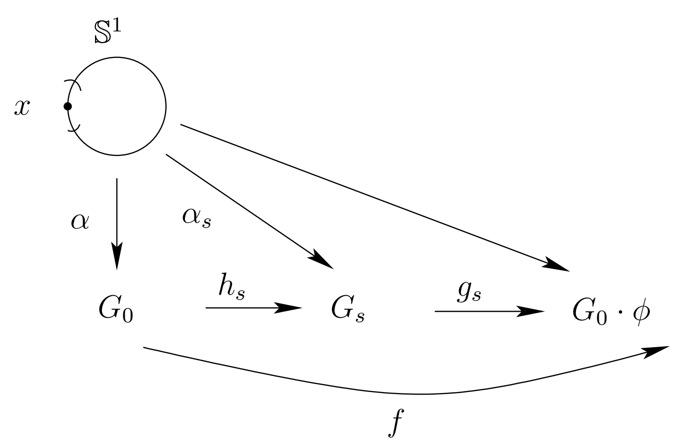

FIGURE 5. If a loop α is illegal with respect to h_s or g_s then it's illegal with respect to f.

$\mathrm{St}_\beta(G_0, G_0\cdot\phi) = \mathrm{St}_\beta(G_0, G_s)\cdot \mathrm{St}_{h(\beta)}(G_s, G_0\cdot\phi)$. Hence

$$d(G_0, G_0\cdot\phi) = d(G_0, G_s) + d(G_s, G_0\cdot\phi)$$

We've shown that $G : [0, \log(\lambda)] \to CV_n$ is a geodesic. We now turn to the claim about the parameterization: let $0 < s < t_1$ then the $d(G_0, G_s)$ is realized by a loop that doesn't contain e_1, e_2 (i.e. any legal loop with respect to h_s) thus $d(G_0, G_s) = \log\left(\frac{e^s l(\alpha, G_0)}{l(\alpha, G_0)}\right) = s$. □

Definition 3.4 (An axis of a fully irreducible automorphism)**.** Let ϕ be a fully irreducible outer automorphism of F_n, $f : G_0 \to G_0$ a train-track representative, and λ the PF eigenvalue of ϕ. Let $G : [0, \log(\lambda)] \to CV_n$ be a (directed) fold line which starts at G_0 and ends at $G_0\cdot\phi$, which is a geodesic parameterized according to arc length. For $t \in \mathbb{R}$ let $k = \left\lfloor \frac{t}{\log\lambda} \right\rfloor$ and define $G(t) = G(t-k)\cdot\phi^k$ (i.e. we translate $G[0, \lambda]$ by ϕ^k and ϕ^{-k}). $\mathcal{L}_f = \mathrm{Im}(G)$ is an invariant geodesic line. We say that $\mathcal{L}_f$ is an axis for ϕ.

Proposition 3.5. *$\mathcal{L}_f$ is a (directed) geodesic parameterized according to arc length.*

Proof. If α is legal at $G(0)$ with respect to $f^2 : G(0) \to G(0)\phi^2$ then α is legal with respect to $f : G(0) \to G(0)\phi$ and $f(\alpha)$ is legal at $G(0)\phi$ with respect to $f : G(0) \to G(0)\phi$. The rest of the argument is as in the proof of proposition 3.3. □

Define $l_\alpha : \mathbb{R} \to \mathbb{R}$ by $l_\alpha(t) = l(\alpha, G(t))$.

3.1. **The projection to an Axis.** Let ϕ be an outer automorphism, and suppose $f : G \to G$ is a stable train-track map for ϕ and $g : H \to H$ is a stable train-track map for ϕ^{-1}. We will show that if α is primitive then there is a bounded set on which $l_\alpha(t)$ achieves its minimum, the bound is uniform over all conjugacy classes α. This will allow us to coarsely define a "nearest point projection" $\pi_f : CV_n \to \mathcal{L}_f$.

We need the following lemmas and notions due to Bestvina, Feighn, and Handel [4].

Definition 3.6. Given a train-track map $f : G \to G$ and a loop α_G in G, the legality of α with respect to the train-track structure of f is

$$LEG_f(\alpha, G) = \frac{\text{Total length of all legal pieces of length } > C_{\text{crit}}}{l(\alpha, G)}$$

Lemma 3.7 (Lemma 5.6 in [4])**.** *If ϕ is non-geometric there is a constant $\epsilon_0 > 0$ and an integer N such that for any conjugacy α:*

$$LEG_f(\phi^N(\alpha), G) > \epsilon_0 \text{ or } LEG_g(\phi^{-N}(\alpha), H) > \epsilon_0$$

We want some version of this lemma that will hold for geometric automorphisms. In the proof of 3.7 the assumption that ϕ is non-geometric is used only to bound the number of consecutive Nielsen paths appearing in α_G. A Nielsen path of f is a path which is periodic under iteration by f. If ϕ is geometric and fully irreducible and f is stable, then there is a unique indivisible Nielsen path β for f in G and this path is in fact a loop. Thus no such bound exists, and indeed the statement in Lemma 3.7 does not hold for β. Therefore, we shall restrict our attention to the case where α is a primitive conjugacy class.

Proposition 3.8. *There is a bound K which depends only on ϕ, such that if α is a basis element then α_G cannot cross more than K consecutive pre-Nielsen paths.*

Proof. The case where ϕ is non-geometric is handled in [4].
If ϕ is geometric then f has a Nielsen loop β. If $f : G \to G$ is stable then β is the only indivisible Nielsen path. Now if α_G crosses two pre-Nielsen loops consecutively then for some m: $f^m(\alpha)$ crosses β twice consecutively. Therefore the Whitehead graph $Wh_G(f^m(\alpha))$ will contain $Wh_G(\beta)$ which is a circle (it corresponds to the boundary of the surface which is a commutator and the whitehead graph of a commutator is a circle). Therefore, $Wh_G(f^m(\alpha))$ will be connected with no cut vertex so by Whitehead's theorem it cannot be a basis element. Thus, if α_G crosses two consecutive pre-Nielsen paths then it cannot be a basis element. □

Putting together proposition 3.8, corollary 2.5 and the proof of 3.7 we get:

Lemma 3.9. *For any irreducible outer automorphism ϕ there is a constant $\epsilon_0 > 0$ and an integer N such that for any primitive conjugacy class α,*

$$LEG_f(\phi^N(\alpha), G) > \epsilon_0 \quad \text{or} \quad LEG_g(\phi^{-N}(\alpha), H) > \epsilon_0$$

if ϕ is non-geometric this holds for all α. In particular, the above holds for a conjugacy class α which corresponds to a candidate α_X in some marked graph X.

Fix a primitive conjugacy class α. Notice that if $LEG_f(\phi^N(\alpha), G) \geq \epsilon$ then $LEG_f(\phi^m(\alpha), G) \geq \epsilon$ for all $m > N$. Define

$$k_0 = \max\{k | LEG_f(\phi^k(\alpha), G) < \epsilon_0\}$$
$$k'_0 = \min\{k | LEG_g(\phi^k(\alpha), H) < \epsilon_0\}$$

Since $LEG_f(\phi^{k_0}(\alpha), G) \not> \epsilon_0$ lemma 3.9 implies $LEG_g(\phi^{k_0-2N}(\alpha), H) > \epsilon_0$ thus

$$|k_0 - k'_0| < 2N_{3.9}$$

The next lemma states that if a loop has legality bounded away from zero, then its growth is exponential.

Lemma 3.10 (Lemma 5.5 in [4]). *For any $A > 0$ there is an N such that if $LEG_f(\alpha, G) > \epsilon_0$ then $l(G, f^N(\alpha)) > A \cdot l(G, \alpha)$.*

We devote the rest of the section to reformulating these results into lemmas that involve the metric.

Corollary 3.11. *There is an $N > 0$ such that for any primitive α both:*

$$l(\phi^{k+N}(\alpha), G) > 2 \cdot l(\phi^k(\alpha), G) \qquad \text{For } k \geq k_0$$
$$l(\phi^{k-N}(\alpha), G) > 2 \cdot l(\phi^k(\alpha), G) \qquad \text{For } k \leq k_0$$

Proof. Let $c_1 = \exp(d(H, G))$ and $c_2 = \exp(d(G, H))$ then for any conjugacy class α:

$$l(\alpha, G) < c_1 l(\alpha, H) \qquad l(\alpha, H) < c_2 l(\alpha, G) \tag{2}$$

Let $N_1 = N_{3.9}$ and $c_3 = \exp(d(H\phi^{-2N_1}, H))$ so for all conjugacy classes α:

$$l(\alpha, H) < c_3 l(\alpha, H\phi^{-2N_1}) \tag{3}$$

In lemma 3.10 take $A = \max\{2, 2c_1c_2c_3\}$ then there is an N_2 such that:

$$\begin{array}{ll} \text{if } LEG_f(\beta, G) > \epsilon_0 \text{ then} & l(\phi^{N_2}(\beta), G) > A \cdot l(\beta, G) \\ \text{if } LEG_g(\beta, H) > \epsilon_0 \text{ then} & l(\phi^{-N_2}(\beta), H) > A \cdot l(\beta, H) \end{array} \tag{4}$$

Let $\beta = \phi^k(\alpha)$, then for $k \geq k_0$: $LEG_f(\phi(\beta), G) > \epsilon_0$ so $l(\phi^{N_2+1}(\beta), G) > 2l(\beta, G)$. For $k \leq k_0$: Recall that $|k_0' - k_0| < 2N_1$ so $LEG_g(\phi^{k_0-2N_1}(\alpha), H) > \epsilon_0$. Therefore, $LEG_g(\phi^{-2N_1}(\beta), H) > \epsilon_0$ and so by equation 4 we get

$$l(\phi^{-2N_2-N_1}(\beta), H) > 2c_1c_2c_3 \cdot l(\phi^{-2N_1}(\beta), H) = \\ 2c_1c_2c_3 \cdot l(\beta, H\phi^{-2N_1}) > 2c_1c_2 \cdot l(\beta, H)$$

Denote $N = 2N_1 + N_2$ then by the previous inequality we get $2c_1 l(\beta, H) < \frac{1}{c_2} l(\phi^{-N}(\beta), H)$. Using this and the formulas in 2 we have

$$2 \cdot l(\beta, G) < 2c_1 \cdot l(\beta, H) < \frac{1}{c_2} \cdot l(\phi^{-N}(\beta), H) < l(\phi^{-N}(\beta), G)$$

□

It follows that for N from Corollary 3.11:

$$l(\phi^{k_0+jN}(\alpha), G) > 2^j \cdot l(\phi^{k_0}(\alpha), G)$$

and

$$l(\phi^{k_0-jN}(\alpha), G) > 2^j \cdot l(\phi^{k_0}(\alpha), G)$$

Let λ be the PF eigenvalue of ϕ and ν the PF eigenvalue for ϕ^{-1}. Let $t_0 = \begin{cases} k_0 \log(\lambda) & \text{if } k_0 > 0 \\ k_0 \log(\mu) & \text{if } k_0 < 0 \end{cases}$ Replacing N with s in the inequalities above, we reformulate them as follows.

Corollary 3.12. *There exists a constant $s > 0$ such that for a primitive α if $|t - t_0| > s$: We define $j = \begin{cases} \lfloor \frac{t-t_0}{s \log \lambda} \rfloor & \text{for } t > t_0 \\ \lfloor \frac{t_0-t}{s \log \mu} \rfloor & \text{for } t < t_0 \end{cases}$ then*

$$\boxed{l(\alpha, G(t)) > 2^j l(\alpha, G(t_0))}$$

Definition 3.13 (min set)**.** For a primitive conjugacy class α: let $L = \min\{l_\alpha(t) \mid t \in \mathbb{R}\}$ and denote by T_α the set of t_α such that $l_\alpha(t_\alpha) = L$. The min set of α is $\pi_f(\alpha) = \{G(t_\alpha) \mid t_\alpha \in T_\alpha\}$.

It follows from corollary 3.12 that $|t_\alpha - t_0| < s_{3.12}$ for every $t_\alpha \in T_\alpha$.

Corollary 3.14. *There is an $s > 0$ such that for every primitive conjugacy class α, $\text{diam}(\pi_f(\alpha)) < s$.*

In subsequent sections we usually suppress the constant s in corollary 3.14. Recall that $|k_0 - k_0'| < 2N_{3.9}$ and that $|t_\alpha - t_0| < s_{3.12}$ thus:

Corollary 3.15. *There is an $s > 0$ such that for every primitive α:*

$$\forall G' \in \pi_f(\alpha) \quad \forall H' \in \pi_g(\alpha): \qquad d(G', H') < s$$

I.e. the min sets of α with respect to $\mathcal{L}_f$ and $\mathcal{L}_g$ are uniformly close.

Corollary 3.16. *There is an $s > 0$ such that for every primitive α, if $t > t_\alpha + s$ then $LEG(\alpha, G(t)) > \epsilon_0$.*

Now we are in a position to define a coarse projection $\pi_f : CV_n \to \mathcal{L}_f$. Let $X \in CV_n$ and $T_X = \{t \mid d(X, G(t)) = d(X, \mathcal{L}_f)\}$. Define the projection of X to $\mathcal{L}_f$: $\pi_f(X) = \{G(t) \mid t \in T_X\}$.

Proposition 3.17. *There is an $s > 0$ such that for every point $X \in CV_n$: $\text{diam}(\pi(X)) < s$.*

Proof. For each candidate α of X: $l_\alpha(t)$ is coarsely decreasing on $(-\infty, t_\alpha - s_{3.12}]$, coarsely increasing on $[t_\alpha + s_{3.12}, \infty)$ and has a minimum in $[t_\alpha - s_{3.12}, t_\alpha + s_{3.12}]$. The function $\mathrm{St}(\alpha_t) = \frac{l_\alpha(t)}{l(\alpha,X)}$ differs from $l_\alpha(t)$ by the multiplicative constant $\frac{1}{l(\alpha,X)}$. So $\mathrm{St}(\alpha_t)$ behaves similarly. Now notice that if $l_1, l_2 : \mathbb{R} \to \mathbb{R}$ are two such functions then $h = \max\{l_1, l_2\}$ also has a coarse minimum see figure 6 . Since there is only a finite number of candidates (depending only on n) then the diameter of $\pi(X)$ is uniformly bounded. □

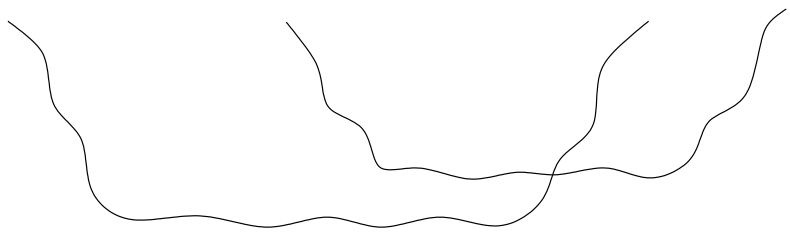

FIGURE 6. If two functions have a coarse minimum then their max has a coarse minimum.

We remark that using the fact that the length map $l(\alpha, t)$ is coarsely exponential we can show that $\pi_f : CV_n \to \mathcal{L}_f$ is coarsely Lipschitz. However we will get a better result in Corollary 5.10.

4. The Whitehead graph of the attracting and repelling laminations

We've defined the Whitehead graph of a conjugacy class α in the basis $\mathcal{B}$. If $X \in CV_n$ is a rose, then $Wh_X(\alpha)$ the Whitehead graph of the loop α is the Whitehead graph of the conjugacy class represented by α in the basis represented by the edges of X. Similarly we can define the Whitehead graph, $Wh_X(\lambda)$ of any quasi-periodic bi-infinite edge path λ in X. This section is devoted to proving lemma 4.2 which produces a rose F in CV_n such that the Whitehead graph of the attracting and repelling laminations of a fully irreducible automorphism does not contain a cut vertex. We remark that the discussion in this section could have been carried out using currents instead of laminations. In that case Lemma 4.1 could be replaced by the stronger results of Kapovich and Lustig in [20].

Let ϕ be a fully irreducible automorphism, and $f : G_0 \to G_0$ a train-track representative for ϕ. Let T_0 be the universal cover of G_0, and $\widetilde{f} : T_0 \to T_0$ a lift of f. Suppose $\Lambda^+_\phi(G_0)$ is the attracting lamination of ϕ realized as bi-infinite lines in G_0. Given a metric tree T in $\mathcal{X}$ (unnormalized outer space) one can define the length of the lamination Λ^+ in T, scaled with respect to T_0, as follows. Let $h : T_0 \to T$ be an equivariant Lipschitz map and let σ be a subsegment of the leaf $\lambda \in \Lambda^+_\phi(G_0)$ and $\widetilde{\sigma}$ be a lift of σ to T_0. Let $[h(\widetilde{\sigma})]$ the tightened image of $\widetilde{\sigma}$ in T then

$$l_{T_0}(\Lambda^+, T) = \lim_{\sigma \to \lambda} \frac{l([h(\widetilde{\sigma})], T)}{l(\widetilde{\sigma}, T_0)} \tag{5}$$

This is not an invariant quantity under resealing the metric of T. Therefore, we modify the definition for $[T]$: Let $[w_1], \ldots, [w_J]$ be a set of conjugacy classes in

F_n which cannot be simultaneously elliptic. Let $\text{tr}(w_1, T), \ldots, \text{tr}(w_J, T)$ be their translation lengths in T and $d(T) = \sum_{i=1}^{J} tr(w_i, T)$ then

$$l_{[T_0]}(\Lambda^+, [T]) = \frac{l_{T_0}(\Lambda^+, T)}{d(T)} \tag{6}$$

Lemma 4.1. *The limit in equation 5 exists, and it is independent of the choice of h. Moreover, the map $l_{[T_0]}(\Lambda^+, \cdot) : \overline{CV_n} \to \mathbb{R}$ is continuous.*

Proof. We begin by showing that the limit exists. This boils down to the fact that λ is quasi-periodic. If $\sigma \subseteq \lambda$ is long enough then σ is a concatenation of a list of words $\tau_1, \ldots, \tau_m$ (like beads on a necklace) which appear with fixed frequencies $r_1, \ldots, r_m$. We can choose the beads/tiles long enough so that the cancellation in $h(\widetilde{\tau_i})h(\widetilde{\tau_j})$ is negligible with respect to the length of τ_i. Thus $h(\widetilde{\sigma})$ (up to small cancellation) is a concatenation of the tiles $h(\widetilde{\tau_i})$, which appear with frequency r_i. So $\frac{l_{T_0}(h(\widetilde{\sigma})),T)}{l_{T_0}(\widetilde{\sigma},T_0)} \sim \frac{\sum_{i=1}^m r_i l(h(\widetilde{\tau_i}),T)}{\sum_{i=1}^m r_i l(\widetilde{\tau_i},T_0)}$. This expression can easily be shown to converge as $k \to \infty$.

Let $L = \text{Lip}(h)$ and $C = BCC(h)$. Denote the edges of G_0 by $e_1, \ldots, e_m$. For each k the i-th k-tile is $\tau_i = f^k(e_i)$ where $1 \le i \le m$. We use $l_i(T_0) = l(\widetilde{\tau_i}, T_0)$, and $l_i(T) = l([h(\widetilde{\tau_i})], T)$ for shorthand but notice that we're suppressing a dependence on k. Let $A = \max\{l_i(T_0) | 1 \le i \le m\}$ and $B = \min\{l_i(T_0) | 1 \le i \le m\}$. Suppose k is large enough so that $\frac{2C}{B} < \epsilon$.

Each leaf λ of Λ_0^+ has a natural 1-tiling by edges in G_0. The standard j-tiling of λ is the f^j image of the 1-tiling of $f^{-j}(\lambda)$. σ is sandwiched between $\sigma_1 \subseteq \sigma_2$ which are both k-tiled and $l(\widetilde{\sigma_1}, T_0) < l(\widetilde{\sigma_2}, T_0) < l(\widetilde{\sigma_1}, T_0) + 2A$.

Let $N_i = \#$occurrences of the tile τ_i in the tiling of σ_1, and $N = \sum_{i=1}^m N_i$. By Perron-Frobenius theory there are $r_1, \ldots, r_m$ such that $\frac{N_i}{N} \to r_i$ as $\sigma_1 \to \lambda$. Let $a_k = \frac{\sum_{i=1}^m r_i l_i(T)}{\sum_{i=1}^m r_i l_i(T_0)}$ where k stands for k-tiles. We show that for large enough σ, $\frac{l(h[\widetilde{\sigma}],T)}{l(\sigma,T_0)}$ is in $[a_k - \epsilon, a_k + \epsilon]$.

We have:

$$\frac{l([h(\widetilde{\sigma}_1)],T)}{l(\widetilde{\sigma}_2,T_0)} \le \frac{l([h(\widetilde{\sigma})],T)}{l(\widetilde{\sigma},T_0)} \le \frac{l([h(\widetilde{\sigma}_2)],T)}{l(\widetilde{\sigma}_1,T_0)}$$

The right hand side limits to:

$$\frac{l([h(\widetilde{\sigma}_2)],T)}{l(\widetilde{\sigma}_1,T_0)} \le \frac{\sum_{i=1}^m N_i l_i(T) + 2AL}{\sum_{i=1}^m N_i l_i(T_0)} =$$

$$\frac{\sum_{i=1}^m \frac{N_i}{N} l_i(T) + \frac{2AL}{N}}{\sum_{i=1}^m \frac{N_i}{N} l_i(T_0)} \xrightarrow[N\to\infty]{} \frac{\sum_{i=1}^m r_i l_i(T)}{\sum_{i=1}^m r_i l_i(T_0)}$$

The left hand side limits to:

$$\begin{aligned} \frac{l_T(h[\widetilde{\sigma}_1])}{l_{T_0}(\widetilde{\sigma}_2)} &\ge \frac{\sum_{i=1}^m N_i [l_i(T) - 2C]}{\sum_{i=1}^m N_i l_i(T_0) + 2A} \\ &= \frac{\sum_{i=1}^m N_i l_i(T)}{\sum_{i=1}^m N_i l_i(T_0) + 2A} - \frac{\sum_{i=1}^m N_i 2C}{\sum_{i=1}^m N_i l_i(T_0) + 2A} \end{aligned}$$

$$\begin{aligned} &\geq \frac{\sum_{i=1}^m \frac{N_i}{N} l_i(T)}{\sum_{i=1}^m \frac{N_i}{N} l_i(T_0)+2A} - \frac{N\cdot 2C}{NB+2A} \\ &\xrightarrow[N\to\infty]{} \frac{\sum_{i=1}^m r_i l_i(T)}{\sum_{i=1}^m r_i l_i(T_0)} - \frac{2C}{B} \\ &\geq \frac{\sum_{i=1}^m r_i l_i(T)}{\sum_{i=1}^m r_i l_i(T_0)} - \epsilon \end{aligned}$$

Thus, for all ϵ, and for large enough σ:

$$a_k - \epsilon \leq \frac{l([h(\widetilde{\sigma})], T)}{l(\widetilde{\sigma}, T_0)} \leq a_k + \epsilon \tag{7}$$

Since $f^N(\tau_i)$ is tiled by τ_i for large enough N, the intervals $[a_{kN} - 2^{-kN}, a_{kN} + 2^{-kN}]$ are nested. By Cantor's nested intervals lemma a_k converges. Thus the limit

$$\lim_{\sigma\to\lambda} \frac{l_T(h[\widetilde{\sigma}])}{l_{T_0}(\widetilde{\sigma})} = c \tag{8}$$

exists.

Next, we show that this limit doesn't depend on the choice of h. We claim that if $h' : T_0 \to T$ is another equivariant Lipschitz map, then $|l([h(\sigma)], T) - l([h'(\sigma)], T)| < 2D$ for some D. Thus the limit in equation 8 is the same for both h and h'. Indeed let p be some point in T_0. Then for all $x \in T_0$ there is a $g \in F_n$ such that $d_{T_0}(x, g\cdot p) \leq 1$. Hence $d(h(x), h'(x)) \leq d(h(x), h(gp)) + d(h(gp), h'(gp)) + d(h'(gp), h'(x)) \leq \mathrm{Lip}(h) + d(h(p), h'(p)) + \mathrm{Lip}(h')$. Denote this constant by D. Thus, for any path $\sigma \subseteq T_0$ the initial and terminal endpoints of $h(\sigma), h'(\sigma)$ are D-close, so $|l([h(\sigma)], T) - l([h'(\sigma)], T)| < 2D$.

Finally we want to show that $l_{[T_0]}(\Lambda^+, [T])$ depends continuously on $[T]$.

$$l_{[T_0]}(\Lambda^+, [T]) = \lim_{k\to\infty} \frac{1}{\sum_{i=1}^m r_i l_i(T_0)} \frac{\sum_{i=1}^m r_i l_i(T)}{d(T)}$$

Without loss of generality suppose $\mathrm{tr}(w_1, T) \neq 0$. If $[T_j] \xrightarrow[j\to\infty]{} [T]$ then $\frac{l_i(T_j)}{\mathrm{tr}(w_1,T_j)} \to \frac{l_i(T)}{\mathrm{tr}(w_1,T)}$ for all $1 \leq i \leq m$ so:

$$\frac{\sum_{i=1}^m r_i l_i(T_j)}{d(T_j)} = \frac{\sum_{i=1}^m r_i l_i(T_j)/\mathrm{tr}(w_1,T_j)}{d(T_j)/\mathrm{tr}(w_1,T_j)} \xrightarrow[j\to\infty]{} \frac{\sum_{i=1}^m r_i l_i(T)/\mathrm{tr}(w_1,T)}{d(T)/\mathrm{tr}(w_1,T)} = \frac{\sum_{i=1}^m r_i l_i(T)}{d(T)}$$

□

Lemma 4.2. *There is a point $F \in CV_n$ such that for any leaves $\lambda \in \Lambda^+_\phi(F)$ and $\nu \in \Lambda^-_\phi(F)$, the whitehead graph $Wh_F(\lambda, \nu)$ is connected and contains no cut vertex.*

To prove this lemma we will need the following proposition proven by Levitt and Lustig [21].

Proposition 4.3. *If $l_{[T_0]}(\Lambda^+, [T]) = 0$ then $l_{[T_0]}(\Lambda^-, [T]) \neq 0$*

Proof. Proposition 5.1 in [21] shows this for a tree T with dense orbits. For a general tree the proof can be found in section 6 of [21]. □

Proof of Lemma 4.2. First recall that if $\lambda_1, \lambda_2 \in \Lambda^+(X)$ are leaves of the attracting lamination then they share the same leaf segments so for any $X \in CV_n$, $Wh_X(\lambda_1) = Wh_X(\lambda_2)$. Since the choice of the leaves doesn't affect the whitehead graph, fix leaves $\lambda \in \Lambda^+$ and $\nu \in \Lambda^-$ once and for all.

Pick a point $X_0 \in CV_n$ whose underlying graph is a rose where all edges have length $\frac{1}{n}$. It was proven in [4] that $Wh_{X_0}(\nu), Wh_{X_0}(\lambda)$ are both connected. If $Wh_{X_0}(\nu) \cup Wh_{X_0}(\lambda)$ contains a cut vertex, then let $X_1 \in CV_n$ be the point obtained from X_0 by leaving the underlying graph unchanged and changing the marking by the whitehead move described right after theorem 1.2. Continue this way to get a sequence $X_0, X_1, X_2, \ldots$ We will show that this process terminates in a finite number of steps with a graph $F = X_N$ for which $Wh_F(\nu) \cup Wh_F(\lambda)$ doesn't contain a cut vertex. A priori, two other cases are possible: $X_k = X_j$ for some $j > k$, and the process never terminates producing an infinite sequence $\{X_i\}_{i=1}^{\infty}$.

Observation 4.4. For all i we have $l_{T_0}(\Lambda^+, \widetilde{X_i}) > l_{T_0}(\Lambda^+, \widetilde{X_{i+1}})$ and $l_{T_0}(\Lambda^-, \widetilde{X_i}) > l_{T_0}(\Lambda^-, \widetilde{X_{i+1}})$

We put off the proof of this observation to show that the Lemma follows. $X_k = X_j$ for $k < j$ is impossible since the lengths get strictly smaller. If the process doesn't terminate then we get an infinite sequence $\{X_i\}_{i=1}^{\infty}$ which has a subsequence converging to $[T] \in \overline{CV_n}$. In CV_n the sequence is discrete, so the limit point must lie in ∂CV_n. We will argue that $l_{[T_0]}(\Lambda^+, [T]) = l_{[T_0]}(\Lambda^-, [T]) = 0$ and get a contradiction to Proposition 4.3.

Let $L = l_{T_0}(\Lambda^+, \widetilde{X_0})$ then $l_{T_0}(\Lambda^+, \widetilde{X_i}) < L$, and together with $d(\widetilde{X_i}) \geq 1$ we get $l_{[T_0]}(\Lambda^+, [\widetilde{X_i}]) < L$. Therefore, $l_{[T_0]}(\Lambda^+, [T]) < L$. Now assume by way of contradiction that $l_{[T_0]}(\Lambda^+, [T]) = L' > 0$. There exists some conjugacy class $[w]$ such that $\mathrm{tr}(w, T) < \frac{L'}{2nL}$ (if T is simplicial then there is a conjugacy class $[w]$ which is elliptic and if T is not simplicial, it has a quotient tree with dense orbits. In either case we can find conjugacy classes with arbitrarily small translation length). Since $\widetilde{X_k}$ converges projectively to $[T]$,

$$\frac{tr(w, \widetilde{X_k})}{d(\widetilde{X_k})} \to \frac{tr(w, T)}{d(T)} < \frac{1}{d(T)} \frac{L'}{2nL} < \frac{L'}{2nL}$$

Thus, for a large enough k, $\frac{\mathrm{tr}(w, \widetilde{X_k})}{d(\widetilde{X_k})} < \frac{1}{nL}$ which implies $\frac{\mathrm{tr}(w, \widetilde{X_k})}{l_{T_0}(\Lambda^+, \widetilde{X_k})} < \frac{1}{nL}$. But this is impossible because $l_{T_0}(\Lambda^+, \widetilde{X_k}) < L$, and $\mathrm{tr}(w, \widetilde{X_k}) > \frac{1}{n}$. So we get a contradiction to $l_{[T_0]}(\Lambda^+, [T]) \neq 0$. A similar argument shows $l_{[T_0]}(\Lambda^-, [T]) = 0$ and we get a contradiction. Therefore, the process must end in a finite number of steps with a graph F such that $Wh_F(\lambda, \nu)$ is connected without a cut vertex. □

Remark 4.5. Experimental evidence suggests that one can actually choose F to lie on an axis of ϕ, but we weren't able to show that.

Proof of Observation 4.4. We must estimate $\lim_{k \to \infty} a_k$ where $a_k = \frac{\sum_{i=1}^m r_i l_i(T)}{\sum_{i=1}^m r_i l_i(T_0)}$ for $T = \widetilde{X}_i$ and $T = \widetilde{X}_{i+1}$. We'll show that lengths of images of tiles in $\widetilde{X_{i+1}}$ are shorter than in $\widetilde{X_i}$. Let $h : G_0 \to X_i$ be a Lipschitz map homotopic to the difference in marking. Let τ be a tile of $\lambda_X \in \Lambda^+_\phi(X)$. The idea is that $[h(\tau)]$ contains a long subsegment of $\lambda_{X_i} = [h(\lambda_X)]$ sandwiched between short segments which don't lie on the leaf since they cancel when we tighten $h(\lambda_X)$. The whitehead move we preform will definitely make $[h(\tau)] \cap \lambda_{X_i}$ shorter, but might possibly make the segments in the beginning and end of $[h(\tau)]$ longer. The following shows that if we take τ to be long enough then the part that gets shorter dominates and the whole tile gets shorter after preforming the Whitehead move.

Let $h : G_0 \to X_i$ be a Lipschitz map homotopic to the difference in marking, $C = BCC(h)$ and $M = \max\{nC, 1\}$. Let e be the oriented edge which represents the cut vertex in $Wh_{X_i}(\lambda, \nu)$. Let X_i^0, X_i^1 be the subgraphs that e cuts off, where $\bar{e} \in X_i^0$. Let $\gamma \subseteq \lambda_{X_i} \in \Lambda^+(X_i)$ be long enough to contain at least $5M$ sequences of the type $\bar{e}x$ and $\bar{x}e$ with $x \in X_i^1$. Let τ_j be tiles in G_0 long enough so that $[h(\tau_j)] \supseteq \gamma$ for all j. We must estimate the difference between $l([h(\tau_j)], X_i)$ and $l([h(\tau_j)], X_{i+1})$. $[h(\tau_j)] = \sigma_1\sigma_2\sigma_3$ where $\sigma_2 \subseteq \lambda_{X_i}$ and $l(\sigma_1, X_i), l(\sigma_3, X_i) < C$ (since the part that isn't in λ_{X_i} is contained in a backtracking segment). Notice that σ_2 only becomes shorter in X_{i+1} but σ_1, σ_3 might become longer. We estimate the growth of σ_1, σ_3. Each of them contains at most nC edges because edge lengths are $\frac{1}{n}$. Each edge might grow by at most $\frac{2}{n}$ (from $\frac{1}{n}$ to $\frac{3}{n}$) so σ_1 contributes at most $nC \cdot \frac{2}{n} = 2C$ additional length. The same is true for σ_3. On the other hand, $\sigma_2 \supseteq \gamma$ and therefore at least $5M$ sequences of the form $\bar{e}x, \bar{x}e$ get substituted by $x, \bar{x}$. Thus the length of σ_2 gets smaller by at least $5M \cdot \frac{1}{n} \geq 5C$. Thus $l(\tau_j, X_{i+1}) - l(\tau_j, X_i) \leq 2 \cdot 2C - 5C < -C$. a_k gets strictly smaller by a definite amount so $l_{T_0}(\Lambda^+, \widetilde{X_i}) > l_{T_0}(\Lambda^+, \widetilde{X_{i+1}})$. By the same argument $l_{T_0}(\Lambda^-, \widetilde{X_i}) > l_{T_0}(\Lambda^-, \widetilde{X_{i+1}})$. □

Remark 4.6. If $Wh_F(\lambda, \nu)$ is connected and does not contain a cut vertex, then $Wh_{F\cdot\phi}(\lambda, \nu)$ and $Wh_{F\cdot\phi^{-1}}(\lambda, \nu)$ satisfy the same property. In fact $Wh_F(\lambda, \nu) = Wh_{F\cdot\phi}(\lambda, \nu) = Wh_{F\cdot\phi^{-1}}(\lambda, \nu)$. Indeed let $k : F \to F$ and $k' : F \to F$ be topological representatives of ϕ, ϕ^{-1} i.e. for each edge they restrict to an immersion or collapse it to a point. Then $Wh_F(\lambda) = Wh_{F\cdot\phi}(k_\#(\lambda))$ this is because $\lambda \in \Lambda^+(F)$ implies $k_\#(\lambda) \in \Lambda^+(F \cdot \phi)$. $Wh_{F\cdot\phi}(k_\#(\lambda)) = Wh_{F\cdot\phi}(\lambda)$ because $\Lambda^+(F \cdot \phi)$ is ϕ, ϕ^{-1}-invariant . Thus $Wh_F(\lambda) = Wh_{F\phi}(\lambda)$. Similarly, $Wh_F(\nu) = Wh_{F\cdot\phi}(k_\#(\nu)) = Wh_{F\cdot\phi}(\nu)$. Thus, $Wh_F(\lambda, \nu) = Wh_{F\cdot\phi}(\lambda, \nu)$. The argument for $Wh_{F\phi^{-1}}(\lambda, \nu)$ is identical.

5. Axes are strongly contracting

Definition 5.1. Let $\Upsilon(X)$ be a lamination in $X \in CV_n$ and η a leaf of $\Upsilon(X)$. Let γ be a tight edge path contained in η. We say that γ is an *r-piece* of η_X if the $l(\gamma, X) \geq r$.

The next proposition states that basis elements can't contain long pieces of both Λ^+ and Λ^-.

Proposition 5.2. *There exists a constant $\ell > 0$ so that for all $G_t \in \mathcal{L}_f$:*

(1) *Let β be a tight loop in G_t representing the conjugacy class $[w]$. Suppose there exist leaves $\lambda \in \Lambda_f^+(G_t)$ and $\nu \in \Lambda_f^-(G_t)$ such that β contains an ℓ-piece of λ and an ℓ-piece of ν. Then w is not a basis element.*
(2) *Let α, β be tight loops in G_t which represent the conjugacy classes $[w_1], [w_2]$, which are compatible with a free decomposition of F_n. If α contains an ℓ-piece of λ (an ℓ-piece of ν) then β doesn't contain an ℓ-piece of ν (an ℓ-piece of λ).*

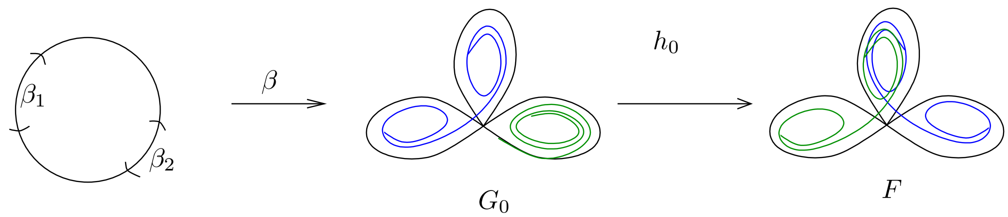

Figure 7. A basis element can't contain long pieces of both laminations

Proof. (1) We first prove this for G_0. By lemma 4.2, there is an $F \in CV_n$ such that $Wh_F(\lambda, \nu)$ is connected and contains no cut point. Suppose $d = d(F, G_0)$ and $k = \exp(d)$ so for all loops α: $\frac{l(\alpha,G_0)}{l(\alpha,F)} \leq k$. Hence $l(\alpha, F) \geq \frac{1}{k} l(\alpha, G_0)$. Let $h_0 : G_0 \to F$ be an optimal Lipschitz map homotopic to the difference in marking. Since λ_F, ν_F are quasi-periodic there is a length r such that if γ_F is an r-piece of λ_F then γ_F contains all of the 2-edge leaf segments in λ_F hence $Wh_F(\lambda_F) = Wh_F(\gamma_F)$. Similarly, if δ_F is an r-piece of ν_F then δ_F contains all of the 2-edge leaf segments in ν_F hence $Wh_F(\nu_F) = Wh_F(\delta_F)$. Let $\ell = k(r + 2C)$ then if $\beta \supseteq \beta_1$ where β_1 is an ℓ-piece of $\lambda_0 \in \Lambda^+(G_0)$ then $l([h_0(\beta_1)], F) > r + 2C$ hence $[h_0(\beta)]$ contains an r-piece of λ_F. Similarly, if $\beta \supseteq \beta_2$ is an ℓ-piece of $\nu_0 \in \Lambda^-(G_0)$ then $l([h_0(\beta_2)], F) > r + 2C$ thus $[h_0(\beta)]$ contains an r-piece of ν_F. Therefore, if β contains such β_1, β_2 (see figure 7) then $Wh_F([h(\beta)]) \supseteq Wh_F(\lambda, \nu)$ and it would be connected and would not contain a cut vertex. By whitehead's theorem $[w]$ is not a basis element.

We can do the same for all graphs $G_t \in \mathcal{L}_f$ and ℓ depends on $d(F, G_t)$, which varies continuously with t. Therefore if we vary t across a fundamental domain of ϕ on $\mathcal{L}_f$, there is an upper bound for ℓ (which we still denote ℓ). Now by remark 4.6 the same is true (with the same ℓ) for any translate of the fundamental domain (we translate F as well so the distance and the optimal map remain the same).

(2) The proof of the second claim is similar to 1. With $Wh_F(\alpha_F, \beta_F)$ replacing $Wh_F(\beta_F)$ in the previous argument.

□

We now turn to prove some applications:

Lemma 5.3. *There is an $s > 0$ such that: if α, β are conjugacy classes which are compatible with a free decomposition of F_n then $|t_\alpha - t_\beta| < s$*

Proof. Denote $t_1 = t_\alpha, t_2 = t_\beta$. Suppose $t_2 > t_1$. Let α_t represent α in G_t, and β_t represent β in G_t. We claim that there is a t_0 such that if $t < t_2 - t_0$ then β_t contains an ℓ-piece of $\nu_{G(t)}$, and if $t > t_1 + t_0$ then α_t contains an ℓ-piece of $\lambda_{G(t)}$. Thus, if $|t_2 - t_1| > 2t_0$ let $r = t_1 + t_0$ then α_r contains an ℓ piece of $\lambda_{G(r)}$ and β_r contains an ℓ piece of $\nu_{G(r)}$ which contradicts proposition 5.2.

To find t_0: by proposition 3.16, there is an $s_1 = s_{3.16}$ such that if $t > t_1 + s_1$ then $LEG_f(\alpha_t, G(t)) > \epsilon_0$. Let $\alpha'_t \subseteq \alpha_t$ be a legal segment of length $> C_{\text{crit}}$. There is an N such that $f^N(\alpha'_t)$ is longer than $\ell + 2BCC(f)$. Let $s_2 = s_1 + N \log(\lambda)$ then at $t_0 = t_1 + s_2$, α contains an ℓ-piece of λ, contributed by α'_t. Similarly for g, the result follows from the fact that $\mathcal{L}_f$ and $\mathcal{L}_g$ are close, and from the fact that t_α and t'_α are close (by corollary 3.15). □

Corollary 5.4. *There exists a constant $s > 0$ such that if α and β are candidates in X then $|t_\alpha - t_\beta| < s$*

Proof. If the candidates have distinct images it follows from corollary 2.5 and by lemma 5.3. If they don't we can apply the previous argument to each of them and a third candidate with an image distinct from theirs to see that $|t_\alpha - t_\beta| < 2s_{5.3}$. □

Corollary 5.5. *There exists an $s > 0$ such that if the translation length of $\alpha \in F_n$ in both X and Y is smaller than 1 then $|\pi(X) - \pi(Y)| < s$.*

Proof. Let $< x_1, x_2, \ldots, x_n >$ be a short basis for $\pi_1(X)$ (all loops are smaller than 1). Since $vol(X) = vol(Y) = 1$, α is carried by a free factor: $< x_1, \ldots, x_k >$. So $|t_\alpha - t_{[x_n]}| < s_{5.3}$. Similarly, for Y,$|t_\alpha - t_{[y_n]}| < s_{5.3}$. So $t_{[x_n]}$ and $t_{[y_n]}$ are uniformly close. By corollary 5.4, we have that t_X and t_Y are uniformly close. □

Consider a geodesic $\mathcal{L}$ in a tree T , and $\pi : T \to \mathcal{L}$ is the closest point projection. The next lemma is motivated by the following observation: If X is a point on $\mathcal{L}$ then $d(Y,X) = d(Y,\pi(Y)) + d(\pi(Y),X)$.

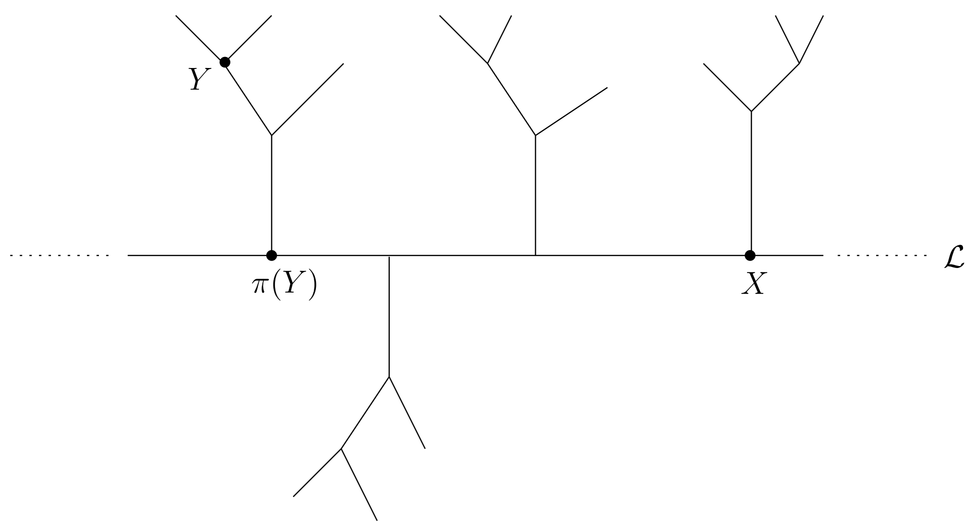

Figure 8. In a tree, the geodesic from Y to a point on a geodesic visits $\pi(Y)$

Lemma 5.6. *There exist constants $s, c > 0$ such that for any Y, if $|t - t_Y| > s$ then $d(Y, G(t)) \geq d(Y,\pi(Y)) + d(\pi(Y), G(t)) - c$*

Proof. Denote $X = G(t)$. Let us first prove it for $t > t_Y$. There is an $s_1 = s_{5.4}$ such that for all candidates α of Y: $|t_\alpha - t_Y| < s_1$. There is an $s_2 = s_{3.16}$ such that if $t > t_\alpha + s_2$ then $LEG_f(\alpha_t, G(t)) > \epsilon_0$. Let $Z = G(t_Y + s_1 + s_2)$ then for any candidate β of Y, $LEG_f(\beta, Z) > \epsilon_0$. Now suppose β_Y in Y is the loop which realizes $d(Y,Z)$, i.e. $\mathrm{St}_\beta(Y,Z) = e^{d(Y,Z)}$. Then, since β is ϵ_0-legal in Z, $l(\beta, X) \geq \epsilon_0 e^{d(Z,X)} l(\beta, Z)$. Thus $\mathrm{St}_\beta(Z,X) \geq \epsilon_0 e^{d(X,Z)}$ so $\mathrm{St}_\beta(Y,X) =$

$\mathrm{St}_\beta(Y,Z)\mathrm{St}_\beta(Z,X) \geq \epsilon_0 e^{d(Y,Z)} e^{d(Z,X)} = \epsilon_0 e^{d(Y,Z)+d(Z,X)}$. We have $\mathrm{St}(Y,X) \geq \mathrm{St}_\beta(Y,X) \geq \epsilon_0 e^{d(Y,Z)+d(Z,X)}$. Thus $d(Y,X) \geq \log(\epsilon_0) + d(Y,Z) + d(Z,X)$. Now recall that $Z = G(t_Y + s_1 + s_2)$ so $d(\pi(Y), Z) = s_1 + s_2$. We have,

$$\begin{aligned} d(Y,Z) &> d(Y,\pi(Y)) \\ d(Z,X) &> d(\pi(Y),X) - (s_1+s_2) \end{aligned}$$

thus $d(Y,X) \geq d(Y,\pi(Y)) + d(\pi(Y),X) - (s_1+s_2) + \log(\epsilon_0)$ let $c = s_1 + s_2 - \log(\epsilon_0)$ and we get $d(Y,X) \geq d(Y,\pi(Y)) + d(\pi(Y),X) - c$.

If $t < t_Y$: there is an s' such that the above holds for g. The claim now follows form the fact that π_f, π_g are uniformly close (see lemma 3.15).

□

Getting back to the tree T, if X, Y are any two points such that $\pi(Y) \neq \pi(X)$ then the geodesic from Y to X passes through $\pi(X)$. In particular $d(Y,X) > d(Y,\pi(X))$. In CV_n:

Lemma 5.7. *There exist constants $s, c > 0$ such that for $X, Y \in CV_n$ if $|t_Y - t_X| > s$, then $d(Y,X) \geq d(Y,\pi(X)) - c$*

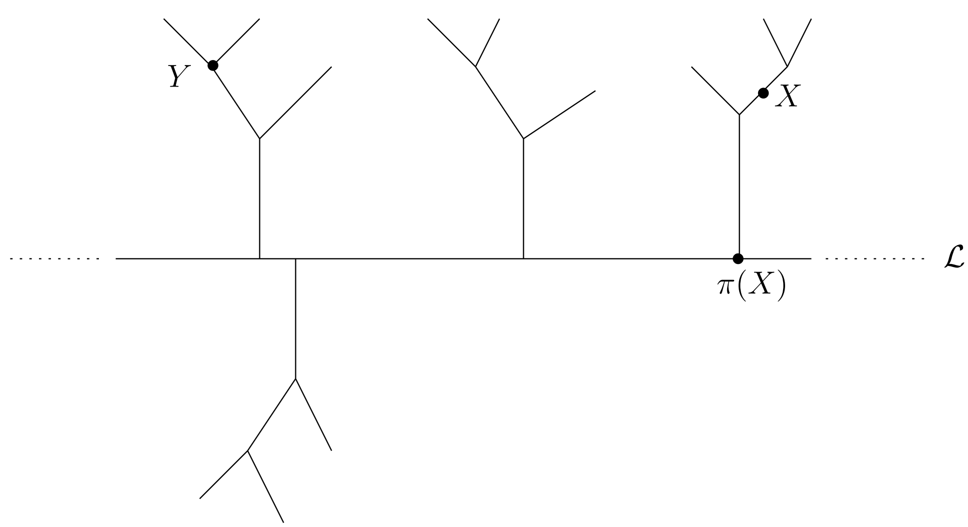

FIGURE 9. In a tree, if X, Y project to different points then the geodesic between them visits both of the projections.

To prove this we recall from lemma 5.2, that if α and β are loops in $G(t)$ representing candidates of X then they cannot contain long pieces of both laminations Λ^+, Λ^-. We will need a slightly souped up version of this.

Proposition 5.8. *Let α_1, α_2 be loops in $G(t) \in \mathcal{L}_f$. Suppose that α_1 is tight as a loop and α_2 tight as a path, and denote by $[\alpha_2]$ the tight loop freely homotopic to α_2. Further assume that the conjugacy classes represented by α_1, α_2 can be completed to a common basis. If $\alpha_1, [\alpha_2]$ both contain an ℓ-piece of $\nu \in \Lambda^-$ then α_2 does not contain an ℓ-piece of $\lambda \in \Lambda^+$.*

Proof. We emphasize that by proposition 5.2, $[\alpha_2]$ doesn't cross an ℓ-piece of λ but we want it not to contain any such pieces in the part that gets cancelled when we tighten the loop.
We represent α_1 by the edge path x in $G(t)$ and α_2 by $u = wyw^{-1}$. We proceed to prove this by way of contradiction. If w crosses an ℓ-piece of λ then $w \nsubseteq x^n$.
Let $x = x_1x_2$ and $w = x^n x_1 w_2$ where $w_2 \neq \emptyset$, $n \geq 0$. There are two cases:

- $x_2 \neq \emptyset$ thus $i(x_2) \neq i(w_2)$. Consider the path
$$ux^{n+1} = x^n x_1 w_2 y w_2^{-1} x_1^{-1} x^{-n} x^{n+1} =$$
$$x^n x_1 w_2 y w_2^{-1} x_1^{-1} x = x^n x_1 w_2 y w_2^{-1} x_2$$
This path is tight since $i(x_2) \neq i(w_2)$. Moreover, it is tight as a loop because x is.
- $x_2 = \emptyset$ then $w = x^n w_2$ where $i(w_2) \neq i(x)$. Again $ux^{n+1} = x^n w_2 y w_2^{-1} x$ is a tight loop.

ux^{n+1} a represents $\alpha_2 \alpha_1^n$ which is a basis element, and it contains an ℓ-piece of ν but since it contains w, it also contains an ℓ-piece of λ. This contradicts proposition 5.2. □

Proof of Lemma 5.7. We prove the claim for X, Y such that $t_Y < t_X$, the case where $t_Y > t_X$ follows by applying the same argument to g. We also make the assumption that X is a bouquet of circles or a graph with two vertices, one separating edge and all other edges are loops.

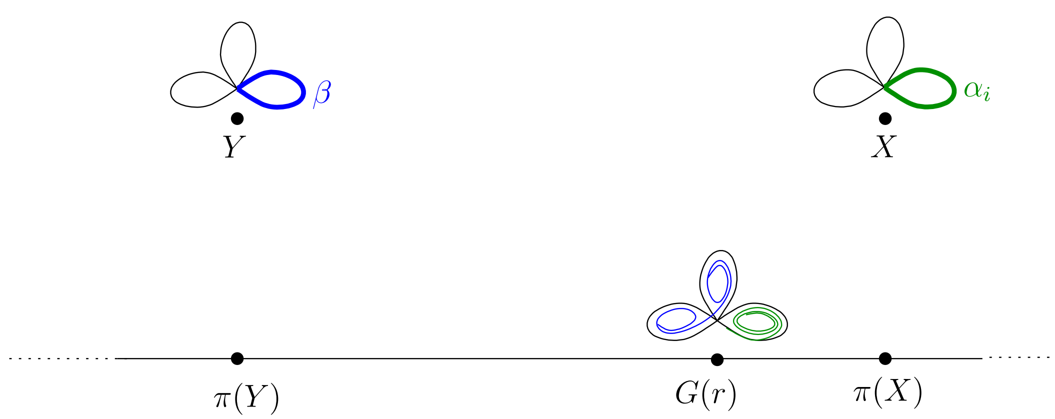

FIGURE 10. In $G(r)$, β contains many l_1-pieces of λ and α_i contain l_1-pieces of ν

The idea is that if $t_Y << t_X$, then for r in the middle of $[t_Y, t_X]$, any loop which is short in Y, would contain many ℓ-pieces of λ in $G(r)$. And any loop which is short in X would contain many ℓ-pieces of ν in $G(r)$, see figure 10. If a candidate in Y was short in X, then it would contain pieces of both λ and ν in $G(r)$ contradicting the fact that it is a basis element. To make the argument precise we need to argue that for a candidate β in Y , $l(\beta, X)$ is longer than a definite fraction of $l(\beta, \pi(X))$. We will show that the number of times β_X crosses any edge of X is bounded below by the number of disjoint ℓ-pieces of λ that appear in $\beta_{\pi(X)}$.

Let $s_1 = s_{5.4}$ i.e. if $t > s_1$ then for any candidate β in Y, $|t_Y - t_\beta| < s_1$. Let $s_2 = s_{3.16}$ i.e. for any primitive conjugacy class β if $t > t_\beta + s_2$ then $LEG_f(\beta, G(t)) > \epsilon_0$. Let s_3 be such that if $t > t_\beta + s_2 + s_3$ then β crosses an ℓ-piece of λ in $G(t)$ (contributed by one of the C_{crit} long legal segments). Let s_4 be such that for any primitive conjugacy class β if $t < t_\beta - s_4$ then β contains an ℓ-piece of ν in $G(t)$. Let $s = 2s_1 + s_2 + s_3 + s_4$ and suppose that $t_X - t_Y > s$ we will show that there exists a c as in the claim.

Let β be a loop in Y such that $d(Y, \pi(X)) = \log(\text{St}_\beta(Y, \pi(X)))$. Then by proposition 5.4 $t_\beta < t_Y + s_1$. Let $r = t_X - s_1 - s_4$ then $r > t_Y + s_1 + s_2 + s_3$. Let $k(r)$ be the number of ℓ-pieces of λ in β_r with disjoint interiors, then

$$k(r) \cdot \ell > \epsilon_0 l(\beta, G(r))$$

If X is a wedge of circles let $\alpha_1, \dots, \alpha_n$ denote the edges in X, and suppose α_1 is the longest edge so that $l(\alpha_1, X) > \frac{1}{n}$. If X is a graph of the other type let α_1 be the longest one-edge-loop, let p be the initial (and terminal) vertex of α_1, and let $\alpha_2, \dots, \alpha_n$ be immersed paths based at p so that the α_is generate $\pi_1(X, p)$. Choose a map $h : X \to G(r)$, homotopic to the difference in marking, so that $h(\alpha_1)$ is a tight loop and $h_\#(\alpha_i)$ are tight as paths. Here, $h_\#(\alpha_i)$ is the immersed path homotopic to $h(\alpha_i)$ relative to $h(p)$. Each $h_\#(\alpha_i)$ in $G(r)$ contains an ℓ-piece of ν. By proposition 5.8 for $1 \le i \le n$, $h_\#(\alpha_i)$ doesn't contain any ℓ-pieces of λ.

Claim. *Let γ be a conjugacy class in F_n and write it as a cyclically reduced word in $\alpha_1, \dots, \alpha_n$ the basis of $\pi_1(X, p)$. If $\gamma_{G(r)}$ contains k occurrences of ℓ-pieces of Λ^+ in $G(r)$ (with disjoint interiors) then γ traverses each α_q at least k times.*

Proof of Claim. First note that if γ_X is a loop which doesn't traverse α_q at all then it is carried by the free factor $< \alpha_1, \dots, \widehat{\alpha_q}, \dots, \alpha_n >$. Using proposition 5.2 applied to $[h(\alpha_q)], [h(\gamma)]$ in $G(r)$, we get that $[h(\gamma)] = \gamma_{G(r)}$ does not contain any l-pieces of Λ^+ in $G(r)$.
Now suppose that $\gamma_X = \alpha_{i_1} \dots \alpha_{i_N}$ so that $\alpha_{i_j} = \alpha_q$ for at most m choices of js. $\gamma_{G(r)}$ is the result of reducing $h_\#(\alpha_{i_1}) \cdot h_\#(\alpha_{i_2}) \cdots h_\#(\alpha_{i_N})$ to get $\sigma_{i_1}\sigma_{i_2} \cdots \sigma_{i_N}$ where σ_{i_j} are paths in $G(r)$ which are what's left from $h_\#(\alpha_{i_j})$ after the reduction (some of which might be trivial).
ℓ-pieces of λ can appear only if they are split between different σ_is. If there were $m+1$ disjoint ℓ-pieces of λ in $\gamma_{G(r)}$ then there is an ℓ-piece of λ appearing in $\sigma_{i_k} \cdots \sigma_{i_l} = [h_\#(\alpha_{i_k}) \dots h_\#(\alpha_{i_l})]$ where none of the α_{i_j} are equal to α_q. This is a contradiction. □

By the claim above β_X in X must traverse α_1 at least $k = k(r)$ times. If $l(\alpha_1, X) > \frac{1}{n+1}$ then $l(\beta, X) > \frac{k}{n+1}$. Otherwise, X has a separating edge e and $l(e, X) > \frac{1}{n+1}$. Let δ be a one-edge-loop so that α_1 and δ are loops on opposite sides of e. By the claim above β_X traverses α_1 and δ alternately at least k times therefore it must cross e at least k times. Again we get $l(\beta, X) > \frac{k}{n+1}$. Therefore, $l(\beta, X) > \frac{\epsilon_0}{(n+1)\ell} l(\beta, G(r))$. Since the distance is almost symmetric, there is a $\mu = \exp\left(\frac{s_1+s_4}{c_{2.10}(\theta)}\right)$ such that $l(\beta, G(r)) > \mu l(\beta, \pi(X))$ therefore $l(\beta, X) > \frac{\epsilon_0 \mu}{(n+1)\ell} l(\beta, \pi(X))$. Thus, we get $\frac{l(\beta,X)}{l(\beta,Y)} > \frac{\epsilon_0\mu}{(n+1)\ell} \frac{l(\beta,\pi(X))}{l(\beta,Y)}$ i.e.

$$d(Y, X) > d(Y, \pi(X)) - \log\left(\frac{(n+1)\ell}{\epsilon_0 \mu}\right)$$

Then c in the statement is the constant $\log\left(\frac{(n+1)\ell}{\epsilon_0\mu}\right)$.

Now we deal with the case that X is not a graph of the types described above. We claim that there is a constant b such that any $X \in CV_n$ is at most a distance b away from a point K whose graph is either a bouquet of circles or K has two vertices connected by a separating edge and all other edges are loops. Moreover, there exists a short loop in X which is still short in K. Therefore, by corollary 5.5 $|\pi(X) - \pi(K)| < s_{5.5}$ so $d(Y,X) > d(Y,K) - d(X,K) \geq d(Y,K) - b > d(Y,\pi(K)) - c - b > d(Y,\pi(X)) - c - b - s_{5.5}$.

To prove that each point in CV_n lies a uniform distance away from a graph K: Let e be the longest edge in X. If e is non-separating let J be a maximal tree in X which doesn't contain e, otherwise let J be the forest obtained from this maximal tree by deleting e. Note that $l(e,X) \geq \frac{1}{3n-3}$. Collapse J to get a new unnormalized graph X' with volume $> \frac{1}{3n-3}$. Notice that X' is a rose if J was a tree otherwise X' is of the second type. Normalize X' to get K. Then $d(X,K) \leq \log\left(\frac{1}{1/(3n-3)}\right) = \log(3n-3)$. The short basis in X is also short in K. □

Corollary 5.9. *There are constants $s, c > 0$ such that:*
If $d(\pi(Y),\pi(X)) > s$ then $d(Y,X) > d(Y,\pi(Y)) + d(\pi(Y),\pi(X)) - c$

Proof. By proposition 5.6 if $d(\pi(Y),\pi(X)) > s_{5.6}$ then

$$d(Y,\pi(X)) > d(Y,\pi(Y)) + d(\pi(Y),\pi(X)) - c_{5.6}$$

By proposition 5.7 if $d(\pi(Y),\pi(X)) > s_{5.7}$ then

$$d(Y,X) > d(Y,\pi(X)) - c_{5.7}$$

So let $s = \max\{s_{5.6}, s_{5.7}\}$ and $c = c_{5.6} + c_{5.7}$ then $d(Y,X) > s$ implies

$$\begin{aligned} d(Y,X) > \quad & d(Y,\pi(X)) - c_{5.7} > d(Y,\pi(Y)) + d(\pi(Y),\pi(X)) - c_{5.6} - c_{5.7} = \\ & d(Y,\pi(Y)) + d(\pi(Y),\pi(X)) - c \end{aligned}$$

□

As a corollary we get that the projection is coarsely Lipschitz.

Corollary 5.10. *There is a constant c such that for all $X, Y \in CV_n$: $d(X,Y) \geq d(\pi(X),\pi(Y)) + c$*

Let $r > 0$. The ball of outward radius r centered at Y is $B_r(Y_{\rightarrow}) = \{X \in CV_n \mid d(Y,X) < r\}$.

Definition 5.11. Let L be a directed geodesic in CV_n, and $\pi_L : CV_n \to L$ is the closest point projection. L is *strongly contracting* if there is a constant $D > 0$ such that for any ball $B \subseteq CV_n$ disjoint from L: $\operatorname{diam}(\pi(B)) < D$.

Strongly contracting geodesics are a feature of negative curvature. If $\mathcal{W}$ is Gromov hyperbolic metric space then all geodesics in $\mathcal{W}$ are uniformly strongly contracting. Moreover, a kind of converse holds as well: if L is a D-strongly contracting geodesic in a metric space $\mathcal{W}$ (not necessarily Gromov hyperbolic), then for any triangle A, B, C with $[B,C]$ contained in L, there are points on $[A,B]$ and $[B,C]$ which are at most a distance $3D$ from $\pi(A)$. If $\mathcal{W}$ also satisfies the fellow traveller property then triangles with one side on L are uniformly slim (in the sense of Gromov). This provides justification for the somewhat informal assertion that $\mathcal{W}$ is negatively curved in the direction of L.

Theorem 5.12. *If $f: G \to G$ is a train-track representative of a fully irreducible outer automorphism ϕ, then $\mathcal{L}_f$ is D-strongly contracting.*

Proof. It is enough to show that there exists a $D > 0$ such that $\text{diam}\{\pi(B_r(Y_\rightarrow))\} < D$ for $r = d(Y, \pi(Y))$. We'll show that if $X \in B_r(Y_\rightarrow)$ then $d(\pi(Y), \pi(X)) < D$ where $D = \max\{s_{5.9}, c_{5.9}\}$. By proposition 5.9 either $d(\pi(Y), \pi(X)) < s_{5.9}$ or $d(Y, X) < d(Y, \pi(Y)) + d(\pi(Y), \pi(X)) - c_{5.9}$. If the latter occurs then

$$\begin{aligned}&r + d(\pi(Y), \pi(X)) - c_{5.9} = \\ &d(Y, \pi(Y)) + d(\pi(Y), \pi(X)) - c_{5.9} < \\ &d(Y, X) < r\end{aligned}$$

Thus $d(\pi(Y), \pi(X)) < c_{5.9}$. □

A well known consequence is:

Lemma 5.13 (The Morse Lemma)**.** *If $\mathcal{L}$ is a D-contracting geodesic and $\mathcal{Q}$ is an (a, b) quasi-geodesic with endpoints on $\mathcal{L}$ then there exists a constant d depending only on D, a, b such that $d_{Haus}(\text{Im}\mathcal{Q}, \mathcal{L}) < d$.*

Remark 5.14. In fact, we only need $\mathcal{Q}$ to satisfy $len(\mathcal{Q}|_{[t_1,t_2]}) < a[d(\mathcal{Q}(t), \mathcal{Q}(t'))] + b$ for the corollary above to hold true. Note that in general, not all quasi-geodesics satisfy this property. However, any quasi-geodesic $\mathcal{Q}$ lies in a bounded Hausdorff neighborhood of a tame quasi-geodesic $\mathcal{Q}'$ therefore we may as well assume $\mathcal{Q}$ satisfies this inequality.

Proof of the Morse Lemma. We fix the following notation. Let $c = \max\{a, b, 1\}$, $R = \max\{d(\mathcal{Q}(t), \mathcal{L}) | t \in \mathbb{R}\}$ and suppose $R > cD$. Let $[s_1, s_2]$ be a maximum subinterval such that for every $s \in [s_1, s_2]$: $d(\mathcal{Q}(s), \mathcal{L}) \geq cD$. Subdivide $[s_1, s_2]$ into: $s_1 = r_1, \ldots, r_m, r_{m+1} = s_2$ where $d(\mathcal{Q}(r_i), \mathcal{Q}(r_{i+1})) = 2cD$ for $i \leq m$ and $d(\mathcal{Q}(r_m), \mathcal{Q}(r_{m+1})) \leq 2cD$. Thus:

$$\text{len}(\mathcal{Q}|_{[s_1,s_2]}) \geq \sum_{i=1}^{m+1} d(\mathcal{Q}(r_i), \mathcal{Q}(r_{i+1})) \geq 2cDm \tag{9}$$

On the other hand, let $P_i = \pi(\mathcal{Q}(r_i))$. Since $d(\mathcal{Q}(r_i), P_i) \geq cD$ we get $d(P_i, P_{i+1}) < D$ therefore $d(P_1, P_{m+1}) \leq D(m+1)$. So $d(\mathcal{Q}(r_1), \mathcal{Q}(r_{m+1})) \leq cD + (m+1)D + c_1cD$ where $c_1 = c_{2.10}(\theta)$ and θ is small enough so that $\mathcal{N}_{cD}(\mathcal{L}) \subseteq CV_n(\theta)$. Therefore,

$$\text{len}(\mathcal{Q}|_{[s_1,s_2]}) \leq cd(\mathcal{Q}(s_1), \mathcal{Q}(s_2)) + c \leq c\big(cD + (m+1)D + c_1cD\big) + c \tag{10}$$

Combining the inequalities 9 and 10 we get:

$$2mcD \leq c^2D + (m+1)cD + c_1c^2D$$

After some manipulation we get: $m \leq \frac{cD(c+c_1)+c}{cD} < 2c + c_1$. Hence $len(\mathcal{Q}|_{[s_1,s_2]}) \leq m2cD < 2(c_1 + 2c)cD$. Thus for each $s \in [s_1, s_2]$:

$$\begin{aligned}d(\mathcal{Q}(s), \mathcal{L}) &< d(\mathcal{Q}(s), \mathcal{Q}(s_2)) + d(\mathcal{Q}(s_2), \mathcal{L}) \leq len(\mathcal{Q}|_{[s_1,s_2]}) + cD < \\ &2(c_1 + 2c)cD + cD\end{aligned}$$

□

6. Axes of fully irreducible elements in the Cayley graph of $Out(F_n)$

Let $\mathscr{C}$ be the Cayley graph of $\mathrm{Out}(F_n)$, with the Whitehead generators $\{\phi_i\}_{i=1}^N$. Let ϕ be a fully irreducible outer automorphism. Let $f : G \to G$ be a train-track map for ϕ. Choose an embedding $\iota : \mathscr{C} \hookrightarrow CV_n$ as follows. Let L be the axis for ϕ in $\mathscr{C}$. Choose some vertex $\psi \in L$ and map $\iota(\psi) = G$. Extend ι to the vertices of $\mathscr{C}$ equivariantly and to the edges of $\mathscr{C}$ by mapping them onto some geodesic between the images of their endpoints. Let $M = \max\{d_{CV_n}(\iota(\phi_i), \iota(\mathrm{id})) \mid \phi_i \text{ is a generator}\}$ then

$$d_{CV_n}(\iota(\psi_1), \iota(\psi_2)) \leq M \cdot d_{\mathscr{C}}(\psi_1, \psi_2)$$

However, this is not a quasi-isometric embedding.

Example 6.1. $\psi_1 = \mathrm{id}$ and $\psi_2 = \begin{cases} x \to x \\ y \to xy^m \end{cases}$ Suppose R is a bouquet of 2 circles each of length $\frac{1}{2}$ with the identity marking. $R\psi_1 = R$, the difference in marking $h : R \to R\psi_2$ is ψ_2 . Thus $d_{CV_n}(R, R\psi_2) = \log(\frac{(m+1)/2}{1/2}) = \log(m+1)$, while $d_{\mathscr{C}}(\psi_1, \psi_2) = m$.

However, for points on L (the axis for ϕ) distances in CV_n coarsely correspond to distances in $\mathscr{C}$. Indeed if α is a legal loop in G then $l(f^m(\alpha), G) = \lambda^m l(\alpha, G)$. Moreover, α in G and $f^m(\alpha)$ in $G \cdot \phi^m$ represent the same conjugacy class, which realizes the maximal stretch. Thus $d_{CV_n}(G, G\phi^m) = m\log(\lambda)$. Let $|\phi|_{\mathscr{C}}$ be the translation length of ϕ, and $\psi \in L$ then $d_{\mathscr{C}}(\psi, \phi^m\psi) = m \cdot |\phi|_{\mathscr{C}} = \frac{d_{CV_n}(G, G\cdot\phi^m)}{\log(\lambda)}|\phi|_{\mathscr{C}} = \frac{|\phi|}{\log(\lambda)} d_{CV_n}(G, G\phi^m) = \frac{|\phi|}{\log(\lambda)} d_{CV_n}(\iota(\psi), \iota(\phi^m\psi))$.

Even though distances in $\mathscr{C}$ are larger (modulo a multiplicative scalar) then their images in CV_n they cannot be arbitrarily larger, as the next lemma shows.

Lemma 6.2. *For every $a > 0$ there is a $b > 0$ such that: If $d_{CV_n}(\iota(\psi), \iota(\chi)) < a$ then $d_{\mathscr{C}}(\psi, \chi) < b$ for all $\psi, \chi \in \mathrm{Out}F_n$.*

Proof. Since the image of ι is discrete, the set $\{\psi \mid d_{CV_n}(\iota(\mathrm{id}), \iota(\psi)) < a\}$ is finite. Let $b = \max\{d_{\mathscr{C}}(\mathrm{id}, \psi) \mid d(\iota(\mathrm{id}), \iota(\psi)) < a\}$. Suppose $d_{CV_n}(\iota(\psi), \iota(\chi)) < a$ then $d_{CV_n}(\iota(\mathrm{id}), \iota(\chi\psi^{-1})) < a$, so $d_{\mathscr{C}}(\mathrm{id}, \chi\psi^{-1}) < b$ and $d_{\mathscr{C}}(\psi, \chi) < b$ □

Theorem 6.3. *L is a stable geodesic in $\mathscr{C}$.*

Proof. Since $d_{CV_n}(\iota(\mathcal{Q}(t)), \iota(\mathcal{Q}(t'))) < M d_{\mathscr{C}}(\mathcal{Q}(t), \mathcal{Q}(t'))$ then the length of $\iota \circ \mathcal{Q}|_{[t,t']}$ in CV_n is smaller or equal to $M\mathrm{len}_{\mathscr{C}}(\mathcal{Q}|_{[t,t']})$. We can assume $\mathcal{Q}$ is a tame quasi-geodesic i.e. $\mathrm{len}_{\mathscr{C}}\mathcal{Q}|_{[t,t']} \leq a'|t-t'| + b'$ where a', b' depend only on a, b (see [9] page 403). Thus $\mathrm{len}(\iota \circ \mathcal{Q}|_{[t,t']}) \leq M(a'|t-t'| + b)$. By remark 5.14, $\mathrm{Im}(\iota \circ \mathcal{Q}) \subseteq \mathcal{N}_d(\mathcal{L}_f)$. By lemma 6.2 we have $\mathrm{Im}(\mathcal{Q}) \subseteq \mathcal{N}_D(L)$ for some D depending only on d. □

7. Applications

Remark 7.1. In this section we change the notation by denoting points in CV_n by lower case letters x, y etc.

7.1. **The asymptotic cone of Outer Space.**

Definition 7.2. A non-principal maximal ultrafilter ω on the integers is a non-empty collection of subsets of $\mathbb{Z}$ which is closed under inclusion, and finite intersection, does not contain any finite sets and if it doesn't contain $A \subset \mathbb{Z}$ then it contains $\mathbb{Z} \smallsetminus A$.
Let ω be a non-principle maximal ultrafilter on the integers. Let (X_i, x_i, d_i) be a sequence of based metric spaces. Define the following pseudo-distance on $\prod_{i\in\mathbb{N}} X_i$:

$$d_\omega(\{a_i\}, \{b_i\}) = \lim_\omega d_{X_i}(a_i, b_i)$$

The ultralimit of (X_i, x_i) is then

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

Where $y \sim y'$ if $d_\omega(y, y') = 0$.
Consider a space X, a point $x \in X$ and a sequence of integers k_i such that $\lim_{i\to\infty} k_i = \infty$. The asymptotic cone of $(X, x, \{k_i\})$ relative to the ultrafilter ω is:

$$\text{Cone}_\omega(X, x, k_i) = \lim_\omega \left(X, x, \frac{1}{k_i} d_X(\cdot, \cdot) \right)$$

The asymptotic cone of a geodesic metric space is a geodesic metric space. Let $p \in CV_n$ be some basepoint, ω a non-principle ultra-filter, and $\{k_i\}$ a sequence of integers with $\lim_{i\to\infty} k_i = \infty$, denote $\mathcal{CV}_n = \text{Cone}_\omega(CV_n, p, \{k_i\})$.

Theorem 7.3. *If $x \in \mathcal{CV}_n$ lies on a geodesic $\mathcal{L} \subseteq \mathcal{CV}_n$ which is the asymptotic cone of an axis $\mathcal{L}_f$ in CV_n of a fully irreducible outer automorphism ϕ, then x is a global cut point of $\mathcal{CV}_n$.*

Proof. We follow the outline of the proof [2] of the analogous fact for $\mathcal{MCG}(S)$. We show that there exists a contraction $p_\mathcal{L} : \mathcal{CV}_n \to \mathcal{L}$ with the following properties:

(1) $p_\mathcal{L}$ restricted to $\mathcal{L}$ is the identity.
(2) $p_\mathcal{L}$ is locally constant on $\mathcal{CV}_n \smallsetminus \mathcal{L}$

We now show how this implies that x is a cut point. Pick a representative $\{x_i\} \in \prod_{i\in\mathbb{N}} CV_n$ of x, we can choose the x_i so they lie on $\mathcal{L}_f$. Let $y_i, z_i \in \mathcal{L}_f$ be points so that x_i is the midpoint of the subsegment of $\mathcal{L}_f$ whose endpoints are y_i, z_i and $d(y_i, x_i) = d(x_i, z_i) = k_i$. Let $y = \{y_i\}, z = \{z_i\}$ then in $\mathcal{CV}_n$: $d(y, x) = 1 = d(x, z)$ and y, z lie on $\mathcal{L}$ with x in between them. Let $\mathcal{H} : [0, 1] \to \mathcal{CV}_n$ be a continuous path from y to z, we claim that it's image must contain the point x. First by connectivity, there is a t such that $p \circ \mathcal{H}(t) = x$. Let $A = (p \circ \mathcal{H})^{-1}(x)$, $t' = \max A$ and since A is closed $t' \in A$. If $\mathcal{H}(t') \notin \mathcal{L}$ then p is locally constant on $\mathcal{H}(t')$ hence t' has a neighborhood contained in A' which is a contradiction. Thus $\mathcal{H}(t') \in \mathcal{L}$ so $H(t') = p \circ \mathcal{H}(t') = x$. This proves that x is a cut point in $\mathcal{CV}_n$.

To finish the proof we show the existence of the map $p_\mathcal{L}$. Let $\{w_i\} \in \prod_{i\in\mathbb{N}} CV_n$ represent a point $w \in \mathcal{CV}_n$. Let $q_i \in \pi_\mathcal{L}(w_i)$. Define $p_\mathcal{L}(w) = \lim_\omega q_i$. If $q_i' \in \pi_\mathcal{L}(w_i)$ are possibly different points, we know that $d(q_i, q_i') < s_{3.17}$ so the ultralimits agree. If $\{w_i'\} \in \prod_{i\in\mathbb{N}} CV_n$ is a different sequence representing w then $\lim_\omega \frac{1}{k_i} d(w_i, w_i') = 0$, but by corollary 5.10 we have $d(\pi(w_i), \pi(w_i')) < d(w_i, w_i') + c_{5.10}$ so the ultralimits of the projections agree. Therefore $p_\mathcal{L}$ is well defined. The fact that $p_\mathcal{L}$ restricts to the identity on $\mathcal{L}$ is now obvious. If $z \in \mathcal{CV}_n \smallsetminus \mathcal{L}$ then let $y \in \mathcal{CV}_n$ be such that $d(z,y) < d(z,p(z))$. Let $\{z_i\},\{y_i\} \in \prod_{i\in\mathbb{N}} CV_n$ represent $z, y \in \mathcal{CV}_n$ then for almost every i we have $\frac{1}{k_i}(d(x_i,q_i) - d(x_i,y_i)) > 0$ hence $d(x_i,y_i) < d(x_i,q_i)$. By theorem 5.12: $d(\pi(x_i),\pi(y_i)) < D_{5.12}$ for almost every i thus $p(x) = p(y)$. □

Behrstock [2] used the hyperbolicity of the curve complex of a surface S of negative Euler characteristic, to show that every point in the asymptotic cone of $\mathcal{MCG}(S)$ is a global cut point. One might hope to use the strongly contracting geodesics in CV_n to prove that the asymptotic cone of $\mathrm{Out}(F_n)$ contains many cut points. One approach might be to map $\mathrm{Cone}_\omega \mathrm{Out}(F_n) \to \mathcal{CV}_n$ and post-compose with $p_{\mathcal{L}}$. We would like to conclude that this map is locally constant off of L_ϕ. However, a point in $\mathrm{Cone}_\omega \mathrm{Out}(F_n) \smallsetminus L_\phi$ might land on $\mathcal{L}_\phi$ in $\mathcal{CV}_n$ thus we cannot immediately deduce that the projection is locally constant off of the geodesic. Nevertheless, we conjecture that these difficulties can be resolved to show that points on axes of fully irreducible automorphisms are cut points in $\mathrm{Cone}_\omega \mathrm{Out}(F_n)$.

We recall the following definition made in [15].

Definition 7.4. Let W be a space with a (possibly non-symmetric) metric and assume W is complete. Let $\mathcal{P}$ be a collection of closed geodesic subsets (called pieces). The space W is said to be tree-graded with respect to $\mathcal{P}$ if the following properties are satisfied:

(1) The intersection of two pieces is either empty of a single point.
(2) Every simple geodesic triangle in X is contained in one piece.

Theorem 7.5. *$\mathcal{CV}_n$ is tree graded.*

Proof. Recall that for each axis $\mathcal{L}_f$ of an irreducible outer automorphism, we defined in the proof of theorem 7.3 a contraction $p_{\mathcal{L}} : \mathcal{CV}_n \to \mathcal{L}$ which was locally constant off of $\mathcal{L}$ and the identity when restricted to $\mathcal{L}$. We redefine $p_{\mathcal{L}} : \mathcal{CV}_n \to \mathbb{R}$ to map onto $\mathbb{R}$ by parameterizing $\mathcal{L}$. Since the set of outer automorphisms is countable, enumerate the axes (one for each fully irreducible outer automorphism) by $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3, \dots$. Define a map $P : \mathcal{CV}_n \to \prod_{i\in\mathbb{N}} \mathbb{R}$ by sending x to the sequence $(p_{\mathcal{L}_i}(x))_{i=1}^{\infty}$. We define the pieces to be pre-images of points in $\prod_{i\in\mathbb{N}} \mathbb{R}$ under P. These are automatically closed and disjoint. We claim that each piece is geodesic. Let $x, y \in \mathcal{CV}_n$ such that $P(x) = P(y)$. Since $\mathcal{CV}_n$ is a geodesic space, there is a geodesic α that connects x to y. We claim that $P \circ \alpha(t) = P(x)$ is constant. If not, then there is a geodesic $\mathcal{L}_i$ such that $p_i \circ \alpha(t) \neq p_i \circ \alpha(0)$. By a similar argument to the one outlined in the proof of theorem 7.3 α must pass through z the midpoint on $\mathcal{L}_i$ between $p_i(\alpha(0))$ and $p_i(\alpha(t))$ (actually it passes through every point on $\mathcal{L}_i$ between these two points). But α must pass through z again on it's way from $\alpha(t)$ to $\alpha(1)$. Thus α contains a loop and so it cannot be a geodesic. Similarly, any loop α such that $P \circ \alpha$ is not constant, must intersect itself and therefore cannot be simple. □

7.2. **Divergence in Outer Space.**

Definition 7.6. Let γ_1, γ_2 be two geodesic rays in CV_n, with $\gamma_1(0) = \gamma_2(0) = x$. The *divergence* function from γ_1 to γ_2 is:

$$\mathrm{div}(\gamma_1,\gamma_2,t) = \inf\left\{ \mathrm{length}(\gamma) \;\middle|\; \begin{array}{c} \gamma : [0,1] \to CV_n \smallsetminus B_r(x_{\leftarrow}) \\ \gamma(0) = \gamma_1(t), \gamma(1) = \gamma_2(t) \end{array} \right\}$$

If $f(t)$ is a function such that:

(1) for every γ_1, γ_2: $\text{div}(\gamma_1, \gamma_2, t) \prec f(t)$ (we use $g(t) \prec f(t)$ to denote the relationship $f(t) \leq c \cdot g(t) + c'$ for all t)
(2) there exist geodesics γ_1, γ_2 such that $\text{div}(\gamma_1, \gamma_2, t) \asymp f(t)$.

then we say that the divergence function of CV_n is on the order of $f(t)$. If only 1 holds we say that f is an upper bound for the divergence of CV_n, if only 2 holds we say that f is a lower bound for the divergence of CV_n.

Duchin and Rafi [16] prove that the divergence in Teichmüller space is quadratic. Behrstock gives an outline of a similar argument [2] for $\mathcal{MCG}(S)$. The proof that the divergence is at least quadratic in the Outer Space setting needs very little modification, but we include it for the reader's convenience.

Proposition 7.7. *Let γ be a path in CV_n, from x to y with $d(\pi(x), \pi(y)) = 2R$. Let z the point on $\mathcal{L}_f$ in the middle of the segment $[\pi(x), \pi(y)] \subseteq \mathcal{L}$. Further assume that the image of γ lies outside the ball $B_R(z_{\leftarrow})$. If $R > 2D_{5.12}$ then there is a constant c such that $Len(\gamma) \geq cR^2$ where c only depends on the constants $D_{5.12}$ and $c_{5.10}$.*

Proof. Subdivide γ into $n > 1$ subsegments $I_1, I_2, \ldots, I_n$, each of which has length $\frac{R}{2}$ except for possibly $Len(I_n) \leq \frac{R}{2}$. Therefore $Len(\gamma) \geq (n-1)\frac{R}{2}$. Let $\mathcal{L}_0$ be the subsegment of $\mathcal{L}_f$ centered at z of length R. Since $\mathcal{L}$ is b-contracting for $b = D_{5.12}$ then $\mathcal{L}_0$ is b'-contracting for $b' = b + 4c_{5.10} + 3$ (see Lemma 3.2 in [7]). Each segment I_j is contained in a ball $B_{R/2}(x'_{\leftarrow})$ disjoint from $\mathcal{L}_0$. Thus the length of each $\pi(I_j) \leq b'$, since these segments cover $\mathcal{L}_0$ we get $R \leq nb'$. Therefore $Len(\gamma) \geq (n-1)\frac{R}{2} > \left(\frac{R}{b'} - 1\right)\frac{R}{2} = \frac{1}{2b'}R^2 - \frac{1}{2}R$. □

The exact behavior of the divergence function of CV_n remains open. Another interesting question is whether the divergence function in $\text{Out}(F_n)$ is quadratic. Behrstock, Druţu and Mosher [1] prove that $\text{Out}(F_n)$ is thick of order at most 1, thus its divergence is at most quadratic.

7.3. **The Behrstock Inequality.** In this section let $\phi, \psi \in \text{Out}(F_n)$ be two irreducible outer automorphisms and f, g their respective train-track representatives. Denote $A = \mathcal{L}_f$, $B = \mathcal{L}_g$ and $p_A = \pi_f$ and $p_B = \pi_g$. Our first goal is to show that either A, B are parallel or the diameter of $p_A(B)$ is bounded, and we'd like to understand what the bound depends on. We introduce the following notation for the next lemma: if $x, y \in A$ denote by $[x, y]_A$ the subinterval of A whose endpoints are x and y.

Lemma 7.8. *There exist constants c, d such that if $x, y \in B$ with $d(p_A(x), p_A(y)) > c$, then*

$$[p_A(x), p_A(y)]_A \subset \mathcal{N}_d(B)$$

c, d depend only on the constants $s_{5.7}, c_{5.7}$ applied to A and B and on $c_{2.10}(\theta)$ where θ is small enough so that $A, B \subseteq CV_n(\theta)$.

Proof. Let $c_1 = c_{2.10}(\theta)$ from corollary 2.10, thus for all $z, w \in CV_n(\theta)$: $\frac{d(z,w)}{c_1} < d(w, z) < c_1 \cdot d(z, w)$. Let s_A, c_A be the constants from lemma 5.7 applied to A, thus if z, w are points such that $d(p_A(z), p_A(w)) > s_A$ then $d(z, w) > d(z, p_A(z)) + d(p_A(z), p_A(w)) - c_A$. Let $a = 1 + (c_1)^2$, $b = c_A(1 + c_1)$ and $d = c_{5.13}(a, b)$ from Corollary 5.13 applied to B, i.e. for every (a, b)-quasi-geodesic $\mathcal{Q}$ with endpoints on B $\mathcal{N}_d(B) \supset \text{Im}\mathcal{Q}$. We prove that $[x, p_A(x)] \cup [p_A(x), p_A(y)]_A \cup [p_A(y), y]$ is an

(a,b)-quasi-geodesic.

First note

$$d(x,y) > d(x,p_A(x)) + d(p_A(x),p_A(y)) - c_A \tag{11}$$

Similarly,

$$\begin{array}{l} d(y,x) > d(y,p_A(y)) + d(p_A(y),p_A(x)) - c_A \\ d(y,x) > d(y,p_A(y)) - c_A > \frac{1}{c_1} d(p_A(y),y) - c_A \end{array}$$

$$(c_1)^2 \cdot d(x,y) > c_1 d(y,x) > d(p_A(y),y) - c_1 c_A \tag{12}$$

Adding equations 11 and 12 we get

$$\begin{array}{l} (1+(c_1)^2) d(x,y) > \\ d(x,p_A(x)) + d(p_A(x),p_A(y)) + d(p_A(y),y) - c_A(1+c_1) \end{array}$$

Therefore $[x,p_A(x)] \cup [p_A(x),p_A(y)]_A \cup [p_A(y),y]$ is a $(1+(c_1)^2, c_A(1+c_1))$-quasi-geodesic. So it is contained in the d neighborhood of B. Hence $[p_A(x),p_A(y)]_A \subset \mathcal{N}_d(B)$. □

The next Lemma is motivated by the following observation. Let X is a proper metric space with a properly discontinuous isometric G-action. Let $g,h \in G$ be hyperbolic isometries of X and let A_g, A_h denote their axes. Then for every d there is a constant k which depends only on $d, |g|, |h|$ such that either A_g, A_h are parallel, or the length of $A_g \cap \mathcal{N}_d(A_h)$ is shorter than k.
In our case, Outer Space is not proper. The example in figure 2 shows that the closure of a ball $B_r(x_{\rightarrow}) = \{y \in CV_n \mid d(x,y) < r\}$ need not be compact. However the closure of the ball $B_r(x_{\leftarrow}) = \{y \in CV_n \mid d(y,x) < r\}$ is always compact. For each $y \in \overline{B_r(x_{\leftarrow})}$ and for all conjugacy classes α, $l(\alpha,y) \geq \frac{l(\alpha,x)}{e^r}$. Thus if θ is the length of the shortest loop in x then $l(\alpha,y) \geq \frac{\theta}{e^r}$. So $\partial CV_n \cap \overline{B_r(x_{\leftarrow})} = \emptyset$ and since $\overline{CV_n}$ is compact then the closure of $B_r(x_{\leftarrow})$ in Outer Space is compact.
Now recall that the $\mathrm{Out}(F_n)$ action is properly discontinuous. Thus for every r there is a number N_r such that $B_r(x_{\leftarrow})$ contains no more than N_r points of any orbit.

Definition 7.9. Let A, B be two axes in $CV_n(\theta)$ and $d > 0$. Let

$$A_B(d) = \{x \in A \mid d(x,p_B(x)) \leq d \text{ and } d(p_B(x),x) \leq d\}$$

Let $convA_B(d)$ be the smallest connected closed set in A containing $A_B(d)$. We claim that for any $a \in convA_B(d)$: $d(a,p_B(a)) < r$ for some r which depends on d and on the constants from Lemma 5.13. The reason is that if $a \in [b,c]_A$ for some $b,c \in A_B$ then $[p_B(b),b]_B \cup [b,c]_A \cup [c,p_B(c)]$ is a $(1,2d)$ quasi-geodesic so it is contained in the r neighborhood of B. Furthermore, $d(p_B(a),a) < c_1 r$ where $c_1 = c_{2.10}(\theta)$.
Let $B_A = \{b \mid b = p_B(x) \text{ for } x \in convA_B\}$ then automatically we have $d(b,p_A(b)) < c_1 r$ therefore $d(p_A(b),b) < (c_1)^2 r$. If $convB_A$ is the smallest closed connected set containing B_A then $\forall b \in convB_A$ we have $d(b,p_A(b)) < R$ for R obtained from Lemma 5.13 applied to A and $(c_1)^2 r$. In conclusion, given d there exists an R depending only on d, c_1 and the constants from Lemma 5.13 such that

$$\begin{array}{llll} \forall x \in convA_B & d(x,p_B(x)) < R & \text{and} & d(p_B(x),x) < R \\ \forall x \in convB_A & d(x,p_A(x)) < R & \text{and} & d(p_A(x),x) < R \end{array}$$

Lemma 7.10. *For every d, there exist constants k, M such that for every f, g such that $A, B \subseteq CV_n(\theta)$ and $|f|, |g| < M$ then either f, g have common powers, or the length of $convA_B(d)$ is smaller than k.*

Proof. Let k denote the length of $convA_B(d)$. Let a be the leftmost point on $convA_B(d)$ assuming that f translates points to the right. Without loss of generality, assume f and g translate points in the same direction. Denote $b = p_B(a)$. For each $i \leq \frac{k}{|f|}$ there is a unique j such that $d(b, p_B(af^i)g^{-j}) < |g|$. Since $d(p_B(af^i), af^i) < R$ then $d(p_B(af^i)g^{-j}, af^ig^{-j}) < R$ hence $d(a, af^ig^{-j}) \leq d(a, b) + d(b, p_B(af^i)g^{-j}) + d(p_B(af^i)g^{-j}, af^ig^{-j}) \leq R + |g| + R = |g| + 2R$. Therefore, $d(af^ig^{-j}, a) < c_1(|g| + 2R) < c_1(M + 2R)$ where $c_1 = c_{2.10}(\theta)$.
Let $r = c_1(M + 2R)$ then there are no more than N_r translates of a in $B_r(a_{\leftarrow})$, but for each $i < \frac{k}{|f|}$: $d(af^ig^{-j}, a) < r$. Therefore, either $k < |f|N_r < MN_r$ or there exists i, j, m, l such that $f^ig^{-j} = f^mg^{-l}$ hence f, g have common powers. □

Corollary 7.11. *There exists a constant k, depending only on the constants from Lemma 7.10 and Lemma 7.8, such that either f, g have common powers or*

$$\mathrm{diam}\{p_A(B)\} < k$$

Proof. Let $\{x_i\}, \{y_i\}$ be sequences on B so that x_i converges to one end of B and y_i to the other. If $d(p_A(x_i), p_A(y_i)) > c_{7.8}$ then $[p_A(x_i), p_A(y_i)] \subseteq \mathcal{N}_d(B)$. Thus $A' = [p_A(x_i), p_A(y_i)]_A$ is contained in $A_B(c_1d)$. Therefore by Lemma 7.10 (with c_1d replacing d) there is a $k_{7.10}$ such that either f, g have common powers or the length of $convA_B(c_1d)$ and hence the length of A' is smaller than $k_{7.10}$. □

Let us go back for a moment to the surface case. We denote by $\mathcal{M}(S)$ the marking complex of S. Let Y, Z be subsurfaces of S, denote by $\mathcal{C}(Z), \mathcal{C}(Y)$ the curve complexes of Z, Y. For definitions of the curve complex and the marking complex consult [2]. Abuse notation to define the projections $p_Y : \mathcal{C}(S) \to \mathcal{C}(Y)$, $p_Y : \mathcal{M}(S) \to \mathcal{C}(Y)$ and $p_Z : \mathcal{C}(S) \to \mathcal{C}(Z)$, $p_Z : \mathcal{M}(S) \to \mathcal{C}(Z)$. In Theorem 4.3 of [2], Behrstock proved that if Y, Z are overlapping subsurfaces of S, neither of which is an annulus, then for any marking μ of S:

$$d_{\mathcal{C}(Y)}(p_Y(\partial Z), p_Y(\mu)) > M \implies d_{\mathcal{C}(Z)}(p_Z(\partial Y), p_Z(\mu)) < M$$

And the constant M depends only on the topological type of S. In other words, if one projection is large then the other must be small. We prove an analogous estimate for our projections.

Suppose f, g, h are train-track maps representing fully irreducible automorphisms and A, B, C are their axes. Suppose that no two of these automorphisms have common powers. We define the coarse distance from B to C with respect to A as $d_A(B, C) = \mathrm{diam}\{p_A(C) \cup p_A(B)\}$.

Lemma 7.12. *There exists a constant $M > 0$ depending only on the constants from Lemma 5.7 and Corollary 7.11 such that at most one of $d_A(B, C), d_B(A, C)$ and $d_C(A, B)$ is greater than M.*

Proof. Let $s_A, c_A, s_B, c_B, s_C, c_C$ be the constants from lemma 5.7 applied to any of the geodesics A, B, C respectively. Let $b \geq c_{7.11}$ the constant from Corollary 7.11 applied to any two of the three geodesics. Let $M > \max\{s_A, c_A, s_B, c_B, s_C, c_C\} +$

2b. We claim that if $d_B(A, C) > M$ then $d_C(A, B) < M$. Assume by way of contradiction that both are greater than M. Let $y \in A$ and $q \in B$ such that $d(y, q) = d_{\text{Haus}}(A, B)$. Let $z = p_C(y) \sim p_C(A)$, $p = p_B(z) \sim p_B(C)$ and $x = p_C(q) \sim p_C(B)$ (see figure 11).

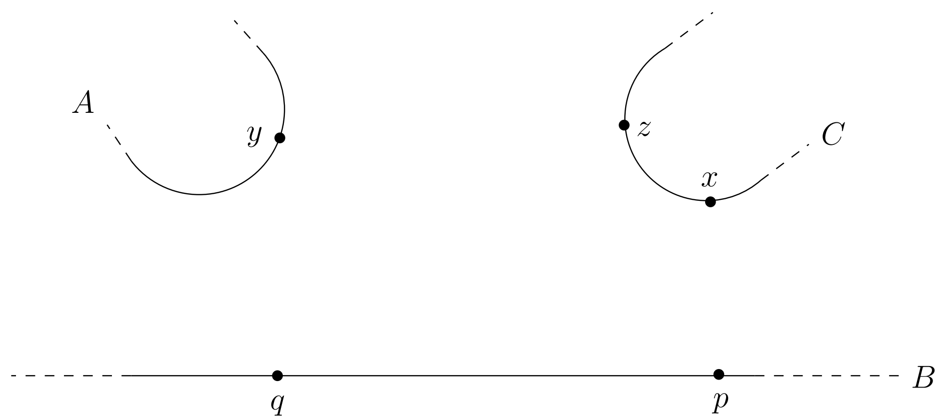

FIGURE 11. If $d(p, q) > M$ then $d(x, z) < M$.

Because $d(p_C(y), p_C(q)) = d(z, x) > M - 2b > s_C$:

$$d(y, q) > d(y, z) + d(z, x) - c_C > d(y, z) + M - 2b - c_C$$

Since $d(p_B(y), p_B(z)) = d(q, p) > M - 2b > s_B$ we have

$$d(y, z) > d(y, q) + d(q, p) - c_B > d(y, q) + M - 2b - c_B$$

Therefore

$$d(y, q) > d(y, z) + M - c_C - 2b > d(y, q) + 2M - c_C - c_B - 4b$$

which implies $2M < c_C + c_B + 4b$ which is a contradiction. □

Theorem 7.13 (The Behrstock inequality)**.** *Let $\phi_1, \ldots, \phi_k$ be fully irreducible outer automorphisms and $f_1, \ldots, f_k$ their respective train track representatives with axes $A_1, \ldots, A_k$. Let $\mathcal{F}$ be the set of translates of $A_1, \ldots, A_k$ under the action of $Out(F_n)$. Then there exists a constant $M > 0$ such that for any $B, C, D \in \mathcal{F}$ if no two of the geodesics are parallel then*

$$d_A(B, C) > M \implies d_B(A, C) < M$$

References

1. Jason Behrstock, Cornelia Druţu, and Lee Mosher, *Thick metric spaces, relative hyperbolicity, and quasi-isometric rigidity*, arXiv:math/0512592v5 math.GT.
2. Jason A. Behrstock, *Asymptotic geometry of the mapping class group and Teichmüller space*, Geom. Topol. **10** (2006), 1523–1578 (electronic). MR **2255505 (2008f:**20108)
3. Mladen Bestvina and Mark Feighn, *Outer limts*, `http://andromeda.rutgers.edu/~feighn/research.html`, 1994, Preprint.
4. Mladen Bestvina, Mark Feighn, and Michael Handel, *Laminations, trees, and irreducible automorphisms of free groups*, Geom. Funct. Anal. **7** (1997), no. 2, 215–244. MR **1445386 (98c:**20045)

5. ______, *The Tits alternative for* Out(F_n)*. I. Dynamics of exponentially-growing automorphisms*, Ann. of Math. (2) **151** (2000), no. 2, 517–623. MR **1765705 (2002a:**20034)
6. ______, *The Tits alternative for* Out(F_n)*. II. A Kolchin type theorem*, Ann. of Math. (2) **161** (2005), no. 1, 1–59. MR **2150382 (2006f:**20030)
7. Mladen Bestvina and Koji Fujiwara, *A characterization of higher rank symmetric spaces via bounded cohomology*, arXiv:math/0702274v3 math.GR.
8. Mladen Bestvina and Michael Handel, *Train tracks and automorphisms of free groups*, Ann. of Math. (2) **135** (1992), no. 1, 1–51. MR **1147956 (92m:**20017)
9. Martin R. Bridson and André Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 319, Springer-Verlag, Berlin, 1999. MR **1744486 (2000k:**53038)
10. Young-Eun Choi and Kasra Rafi, *Comparison between Teichmüller and Lipschitz metrics*, J. Lond. Math. Soc. (2) **76** (2007), no. 3, 739–756. MR 2377122
11. Marshall M. Cohen and Martin Lustig, *Very small group actions on* **R***-trees and Dehn twist automorphisms*, Topology **34** (1995), no. 3, 575–617. MR **1341810 (96g:**20053)
12. Daryl Cooper, *Automorphisms of free groups have finitely generated fixed point sets*, J. Algebra **111** (1987), no. 2, 453–456. MR **916179 (89a:**20024)
13. Marc Culler and John W. Morgan, *Group actions on* **R***-trees*, Proc. London Math. Soc. (3) **55** (1987), no. 3, 571–604. MR **907233 (88f:**20055)
14. Marc Culler and Karen Vogtmann, *Moduli of graphs and automorphisms of free groups*, Invent. Math. **84** (1986), no. 1, 91–119. MR **830040 (87f:**20048)
15. Cornelia Druţu and Mark Sapir, *Tree-graded spaces and asymptotic cones of groups*, Topology **44** (2005), no. 5, 959–1058, With an appendix by Denis Osin and Sapir. MR **2153979 (2006d:**20078)
16. Moon Duchin and Kasra Rafi, *Divergence of geodesics in teichmuller space and the mapping class group*, arXiv:math/0611359v2 math.GT.
17. Stefano Francaviglia and Armando Martino, *Metric properties of outer space*, arXiv:0803.0640v1 math.GR.
18. Michael Handel and Lee Mosher, *Axes in outer space*, arXiv:math/0605355v2 math.GR.
19. ______, *Parageometric outer automorphisms of free groups*, Trans. Amer. Math. Soc. **359** (2007), no. 7, 3153–3183 (electronic). MR **2299450 (2008c:**20045)
20. Ilya Kapovich and Martin Lustig, *Geometric intersection number and analogues of the curve complex for free groups*, arXiv:0711.3806 math.GR.
21. Gilbert Levitt and Martin Lustig, *Irreducible automorphisms of* F_n *have north-south dynamics on compactified outer space*, J. Inst. Math. Jussieu **2** (2003), no. 1, 59–72. MR **1955207 (2004a:**20046)
22. Reiner Martin, *Non-uniquely ergodic foliations of thin type, measured currents, and automorphisms of free groups*, Ph.D. thesis, UCLA, 1995.
23. Howard Masur, *On a class of geodesics in Teichmüller space*, Ann. of Math. (2) **102** (1975), no. 2, 205–221. MR 0385173 (52 #6038)
24. Howard A. Masur and Michael Wolf, *Teichmüller space is not Gromov hyperbolic*, Ann. Acad. Sci. Fenn. Ser. A I Math. **20** (1995), no. 2, 259–267. MR **1346811 (96f:**30048)
25. Yair N. Minsky, *Quasi-projections in Teichmüller space*, J. Reine Angew. Math. **473** (1996), 121–136. MR **1390685 (97b:**32020)
26. Frédéric Paulin, *The Gromov topology on* **R***-trees*, Topology Appl. **32** (1989), no. 3, 197–221. MR **1007101 (90k:**57015)
27. William Thurston, *Minimal stretch maps between hyperbolic surfaces*, arXiv:math.GT/9801039 math.GT.
28. J. H. C. Whitehead, *On certain sets of elements in a free group*, Proc. London Math. Soc. **41** (1936), 48–56.