The co-surface graph and the geometry of hyperbolic free group extensions

Spencer Dowdall and Samuel J. Taylor [*]

November 28, 2016

Abstract

We introduce the co-surface graph $\mathcal{CS}$ of a finitely generated free group $\mathbb{F}$ and use it to study the geometry of hyperbolic group extensions of $\mathbb{F}$. Among other things, we show that the Gromov boundary of the co-surface graph is equivariantly homeomorphic to the space of free arational $\mathbb{F}$–trees and use this to prove that a finitely generated subgroup of $\mathrm{Out}(\mathbb{F})$ quasi-isometrically embeds into the co-surface graph if and only if it is purely atoroidal and quasi-isometrically embeds into the free factor complex. This answers a question of I. Kapovich. Our earlier work [DT] shows that every such group gives rise to a hyperbolic extension of $\mathbb{F}$, and here we prove a converse to this result that characterizes the hyperbolic extensions of $\mathbb{F}$ arising in this manner. As an application of our techniques, we additionally obtain a Scott–Swarup type theorem for this class of extensions.

1 Introduction

Let $\mathbb{F}$ be the free group of rank $r \geq 3$ and let $\mathrm{Out}(\mathbb{F})$ be its outer automorphism group. Every subgroup $\Gamma \leq \mathrm{Out}(\mathbb{F})$ gives rise to an exact sequence

$$1 \longrightarrow \mathbb{F} \xrightarrow{\ i\ } E_\Gamma \xrightarrow{\ p\ } \Gamma \longrightarrow 1, \tag{1}$$

in which E_Γ is the preimage of Γ under the homomorphism $\mathrm{Aut}(\mathbb{F}) \to \mathrm{Out}(\mathbb{F})$ and $\mathbb{F} \lhd E_\Gamma$ is identified with the inner automorphisms $\mathrm{Inn}(\mathbb{F})$. In fact every group extension of $\mathbb{F}$ surjects onto an extension E_Γ of this form. In [DT], we gave conditions on $\Gamma \leq \mathrm{Out}(\mathbb{F})$ that guarantee the associated extension E_Γ is Gromov hyperbolic. To state these conditions, first recall that Γ is purely atoroidal if each infinite order element is atoroidal (no power fixes a nontrivial conjugacy class of $\mathbb{F}$) and that the free factor complex $\mathcal{F}$ is the simplicial complex whose k–simplices are nested chains $A_0 < \cdots < A_k$ of proper free factors of $\mathbb{F}$, up to conjugacy. Outer automorphisms act isometrically on $\mathcal{F}$, and we say that a finitely generated subgroup $\Gamma \leq \mathrm{Out}(\mathbb{F})$ **qi-embeds** into $\mathcal{F}$ if some (equivalently any) orbit map $\Gamma \to \mathcal{F}$ is a quasi-isometric embedding.

Theorem 1.1 ([DT])**.** *Suppose that a finitely generated subgroup $\Gamma \leq \mathrm{Out}(\mathbb{F})$ is purely atoroidal and qi-embeds into $\mathcal{F}$. Then the free group extension E_Γ in Equation (1) is hyperbolic.*

The goal of the present paper is twofold: to refine these conditions for hyperbolicity, and to make a more in depth study of the geometry of these hyperbolic extensions. This study culminates in a converse to the above result that characterizes those hyperbolic extensions arising from Theorem 1.1. We note that the exact converse of Theorem 1.1 is well-known to be false.

Following Hamenstädt and Hensel [HH], a subgroup $\Gamma \leq \text{Out}(\mathbb{F})$ is said to be **convex cocompact** if it qi-embeds into $\mathcal{F}$. However, unlike the analogous situation for mapping class groups [KL3, Ham1], convex cocompactness itself does not ensure hyperbolicity of E_Γ. Indeed, pure atoroidality of Γ is essential for E_Γ to be hyperbolic, since a periodic conjugacy class for $\phi \in \Gamma$ gives rise to a $\mathbb{Z} \oplus \mathbb{Z}$ in E_Γ. Further, there are automorphisms of $\mathbb{F}$ that act loxodromically on $\mathcal{F}$ but are not atoroidal. In fact, by combining work of Bestvina–Handel [BH1] and Bestvina–Feighn [BF2], such automorphisms precisely correspond to pseudo-Anosov homeomorphisms of once-punctured, possibly nonorientable, surfaces.

This suggests that $\mathcal{F}$ is not the correct complex for studying hyperbolic extensions of $\mathbb{F}$. It is natural to build a better-suited complex by starting with $\mathcal{F}$ and coning off the curve graphs for all once-punctured surfaces S with $\pi_1(S) \cong \mathbb{F}$. Versions of this construction appear several places in the literature—first in the work of Kapovich–Lustig [KL1] and later Mann–Reynolds [MR1] and Mann [Man] (see §4)—in each case producing a hyperbolic $\text{Out}(\mathbb{F})$–graph Y with the property that any subgroup Γ that qi-embeds into Y is both convex cocompact and purely atoroidal. The converse was posed as a question by I. Kapovich:

Question 1.2 (I. Kapovich). Suppose that $\Gamma \leq \text{Out}(\mathbb{F})$ is purely atoroidal and convex compact. Is the orbit map $\Gamma \to Y$ a quasi-isomeric embedding?

To answer Question 1.2, we introduce (§4) a new model for the graph Y that is both simple to define and natural for our purposes. This **co-surface graph** is defined to be the simplicial graph $\mathcal{CS}$ whose vertices are conjugacy classes of primitive elements of $\mathbb{F}$ and where two conjugacy classes are adjacent if there is a once-punctured surface S with $\pi_1(S) \cong \mathbb{F}$ in which they are both represented by simple closed curves on S.

Theorem 4.12 (Qi-embedding into $\mathcal{CS}$). *Let Γ be finitely generated subgroup of $\text{Out}(\mathbb{F})$. Then Γ qi-embeds into the co-surface graph $\mathcal{CS}$ if and only if Γ is purely atoroidal and convex cocompact.*

After formulating Question 1.2, Kapovich showed it cannot be answered from formal properties of the action $\Gamma \curvearrowright Y$. That is, Kapovich constructs an action of the free group of rank 2 on a hyperbolic graph X which has all the properties of the action $\Gamma \curvearrowright Y$ but whose orbit map $\Gamma \to X$ is not a qi-embedding [Kap2]. Thus the proof of Theorem 4.12 necessarily requires a deeper understanding of the co-surface graph itself. Indeed, our argument uses the fine geometric structure of Culler and Vogtmann's Outer space $\mathcal{X}$ and the following calculation of the Gromov boundary of $\mathcal{CS}$:

Theorem 4.8 (Boundary of $\mathcal{CS}$). *The Gromov boundary $\partial \mathcal{CS}$ of the co-surface graph is $\text{Out}(\mathbb{F})$–equivariantly homeomorphic to the subspace of $\partial \mathcal{F}$ consisting of classes of free arational trees.*

We obtain Theorem 4.8 as a corollary of the general theory of **alignment-preserving maps** that we develop in §3 and which may be of independent interest. Briefly, three (ordered) points are coarsely aligned if the triangle inequality for them is nearly an equality, and a map that respects this condition is said to be alignment preserving. We show (Theorem 3.2) that any coarsely surjective alignment preserving map $X \to Y$ between hyperbolic metric spaces extends to a homeomorphism between ∂Y and a specific subset of ∂X.

The co-surface graph $\mathcal{CS}$ has other advantages over the factor complex $\mathcal{F}$, and indeed this is a major theme of the present paper. For example, it is well known (see [BFH]) that full irreducibility is not stable under passage to finite index subgroups. This causes complications when attempting to study the subgroup structure of E_Γ. However, the following result shows that this is not an issue for $\mathcal{CS}$:

Proposition 5.2. *Let H be a finite index subgroup of $\mathbb{F}$ and let Γ^H denote the subgroup of $\text{Out}(H)$ induced by elements of $\Gamma \leq \text{Out}(\mathbb{F})$ that stabilize the conjugacy class of H. If Γ is finitely generated and qi-embeds into $\mathcal{CS}$, then Γ^H also qi-embeds into $\mathcal{CS}(H)$.*

Proposition 5.2 is one of the key ingredients allowing us to establish a Scott–Swarup [SS] type theorem for these extensions of $\mathbb{F}$. Recall that the fiber subgroup $\mathbb{F}$, being infinite and normal, is exponentially distorted in the hyperbolic group E_Γ. The following theorem, however, shows that such distortion is confined to finite index subgroups of $\mathbb{F}$; this mirrors a result of Dowdall–Kent–Leininger [DKL] for hyperbolic surface group extensions. Bear in mind that the statement is false without the hypothesis that Γ qi-embed into $\mathcal{CS}$ (see §7.4).

Theorem 7.9 (Nondistortion in fibers). *Suppose that $\Gamma \leq \mathrm{Out}(\mathbb{F})$ quasi-isometrically embeds into $\mathcal{CS}$, and let L be a finitely generated subgroup of the fiber $\mathbb{F} \lhd E_\Gamma$. Then L is quasiconvex, and hence undistorted, in the hyperbolic extension E_Γ if and only if L has infinite index in $\mathbb{F}$.*

We note that Mj and Rafi [MR2] have recently, and independently, proven Theorem 7.9 by very different methods. Their approach uses structural results on convex cocompact subgroups proven in [DT] as well as a characterization of the Cannon–Thurston map for (1) that we obtained with Kapovich in [DKT] and which builds on earlier work of Mj [Mit1]. Our proof is more direct and proceeds as follows.

The second key ingredient needed to prove Theorem 7.9 is a careful study of the geometry of hyperbolic extensions E_Γ that focuses on the relationship between the "local" axis of an element $a \in \mathbb{F}$ acting on any given fiber of $p\colon E_\Gamma \to \Gamma$ and the "global" axis for a acting on E_Γ. Specifically, if a^* denotes the geodesic in E_Γ whose endpoints are the fixed points $a^{\pm\infty}$ in ∂E_Γ, we define the **width** of a to be the quantity

$$\mathrm{width}(a) = \mathrm{diam}_\Gamma \, p(a^*).$$

This concept was first studied in the context of surface group extensions by Kent and Leininger [KL4]. We prove (Theorem 7.2) that when Γ is convex cocompact, the quantity $\mathrm{width}(a)$ is uniformly bounded over all simple elements $a \in \mathbb{F}$, where an element is **simple** if it is contained in some proper free factor of $\mathbb{F}$. As a consequence, we show that the global axis a^* fellow travels the local axis for $a \in \mathbb{F}$ acting on the fiber of $E_\Gamma \to \Gamma$ in which the translation length of a is minimized. Combining this with Proposition 5.2 leads to Theorem 7.9. This analysis also allows us to prove the following theorem, which gives a converse to Theorem 1.1 and characterizes hyperbolic extensions arising from convex cocompact subgroups as those for which the simple elements have uniformly bounded width:

Theorem 8.1 (Convex cocompactness). *Suppose that $1 \to \mathbb{F} \to E \to Q \to 1$ is a hyperbolic extension of $\mathbb{F}$. Then Q has convex cocompact image in $\mathrm{Out}(\mathbb{F})$ (and hence admits a quasi-isometric embedding orbit map into $\mathcal{CS}$) if and only there exists $D \geq 0$ so that $\mathrm{width}_Q(a) \leq D$ for each simple element $a \in \mathbb{F}$.*

2 Background

Throughout, $\mathbb{F}$ will denote a finitely generated free group of rank $r = \mathrm{rk}(\mathbb{F})$ at least 3. In this section we review several structures associated to $\mathbb{F}$ that will be relevant to our work.

2.1 Coarse geometry

A map $f\colon X \to Y$ of metric spaces is a K**–quasi-isometric embedding** if

$$d_X(a,b)/K - K \leq d_Y(f(a),f(b)) \leq Kd_X(a,b) + K$$

for all $a,b \in X$. The map is moreover a K**–quasi-isometry** if its image is K–dense in Y. A K**–quasigeodesic** is then a K–quasi-isometric embedding of an interval $\mathbf{I} \subset \mathbb{R}$ into a metric space. The **Hausdorff distance** between two subsets A, B of a metric space X is the infimum of all $\varepsilon > 0$ for which A and B are both contained within the ε–neighborhoods of each other.

A geodesic metric space X is δ**–hyperbolic**, where $\delta \geq 0$, if every geodesic triangle in X is δ–thin, meaning that each side is contained within the δ–neighborhood of the other two. Every such space has a well defined **Gromov boundary** ∂X consisting of equivalence classes of admissible sequences in X, where a sequence $\{a_n\}$ is admissible if $\lim_{n,m}(a_n|a_m)_x = \infty$ and two sequence $\{a_n\}$ and $\{b_n\}$ are equivalent if

$\lim_{n,m}(a_n|b_m)_x = \infty$ for some $x \in X$. Here, $(a|b)_x$ denotes the **Gromov product** $(d(a,x)+d(b,x)-d(a,b))/2$. One says the admissible sequence $a_1, a_2, \ldots \in X$ **converges** to the point $\{a_n\} \in \partial X$. The Gromov product, with respect to $x \in X$, may be extended to ∂X by declaring

$$(a|b)_x := \sup \liminf_{m,n \to \infty} (a_m|b_n)_x,$$

where the supremum here is over all sequences $\{a_m\}, \{b_n\}$ converging respectively to $a, b \in \partial X$. The boundary ∂X is then equipped with the topology in which the sets $N_a^x(r) := \{b \in \partial X : (a|b)_x \geq r\}$ give a basis of open neighborhoods about the point $a \in \partial X$. Moreover, the topologies on X and ∂X may be extended to a topology on the disjoint union $X \cup \partial X$ for which a sequence $x_n \in X$ converges to $\zeta \in \partial X$ if and only if $\{x_n\}$ is admissible and equivalent to ζ. When X is **proper** (meaning that closed metric balls are compact), $X \cup \partial X$ is a compactification of X. Finally, hyperbolicity itself and the Gromov boundary ∂X are both quasi-isometry invariants of X. See [BH2, Section III.H.3], [GdlH], or [KB, Section 2] for more details.

If X is δ–hyperbolic, then every quasigeodesic ray $r \colon [0,\infty) \to X$ converges to its **endpoint at infinity** $r(\infty) := \{r(n)\}_{n=1}^\infty \in \partial X$, and any two rays whose images have finite Hausdorff distance determine the same endpoint. Conversely, as explained in [KB, Remark 2.16], for any $x_0 \in X$ and $\zeta \in \partial X$ one may build a 10δ quasigeodesic $r \colon \mathbb{R}_+ \to X$ with the properties that $r(0) = x_0$ and $r(\infty) = \zeta$.

Throughout, we will use $\mathbf{I}$ (and sometimes $\mathbf{J}$) to denote a closed subinterval of $\mathbb{R}$. We write $\mathbf{I}_-$ and $\mathbf{I}_+$ for the infimum and supremum of $\mathbf{I}$, respectively. With this notation, every quasigeodesic $\gamma \colon \mathbf{I} \to X$ naturally has two well-defined **endpoints** $\gamma(\mathbf{I}_\pm) \in X \cup \partial X$, where $\gamma(\mathbf{I}_\pm) \in \partial X$ if $\mathbf{I}_\pm = \pm\infty$ and $\gamma(\mathbf{I}_\pm) \in X$ otherwise. The following is a fundamental feature of hyperbolic metric spaces; see [BH2, Theorem III.H.1.7] for a proof.

Proposition 2.1 (Stability of quasigeodesics)**.** *For every $K \geq 1$ and $\delta \geq 0$ there exists a **stability constant** $R_0(K,\delta) > 0$ such that if $\gamma \colon \mathbf{I} \to X$ and $\rho \colon \mathbf{J} \to X$ are K–quasigeodesics with the same endpoints in $X \cup \partial X$ for a δ–hyperbolic space X, then $\gamma(\mathbf{I})$ and $\rho(\mathbf{J})$ have Hausdorff distance at most $R_0(K,\delta)$.*

2.2 Currents and laminations

All finite-valent Cayley graphs of $\mathbb{F}$ are quasi-isometric hyperbolic spaces, and we write $\partial\mathbb{F}$ to denote their common Gromov boundary. The free group $\mathbb{F}$ acts on each of its Cayley graphs by left multiplication, and this extends to a left action of $\mathbb{F}$ on $\partial\mathbb{F}$ by homeomorphisms. Let $\partial^2\mathbb{F} = \{(\eta,\xi) \mid \eta,\xi \in \partial\mathbb{F}, \eta \neq \xi\}$ denote the double boundary of $\mathbb{F}$, equipped with the subspace topology from $\partial\mathbb{F} \times \partial\mathbb{F}$. A **lamination** on $\mathbb{F}$ is a nonempty closed subset of $\partial^2\mathbb{F}$ that is invariant under both the flip map $(\eta,\xi) \mapsto (\xi,\eta)$ and the (diagonal) action of $\mathbb{F}$. A lamination is **minimal** if it does not contain a proper sublamination. If L is a lamination on $\mathbb{F}$, we write L' to denote the set of accumulation points of L in $\partial^2\mathbb{F}$. Note that $L' \subset L$, since L is closed, and that L' is itself a lamination.

Following [Kap1], a **geodesic current** on $\mathbb{F}$ is a positive Radon measure on $\partial^2\mathbb{F}$ that is both flip and $\mathbb{F}$–invariant. Notice that the support $\mathrm{supp}(\mu)$ of every nonzero current μ is necessarily a lamination. We write $\mathrm{Curr}(\mathbb{F})$ for the space of all geodesic currents on $\mathbb{F}$ equipped with the weak topology. Quotienting by the action of $\mathbb{R}_+$ by scalar multiplication yields the compact space $\mathbb{P}\mathrm{Curr}(\mathbb{F})$ of **projective geodesic currents**.

Let us discuss some basic examples of these concepts. Every nontrivial $a \in \mathbb{F}$ acts on $\partial\mathbb{F}$ with a unique attracting fixed point a^+ and repelling fixed point a^-. The **lamination of a (nontrivial) conjugacy class** α of $\mathbb{F}$ is then defined to be

$$L(\alpha) := \bigcup_{a \in \alpha} \{(a^+,a^-),(a^-,a^+)\};$$

notice that $L(\alpha)$ is indeed a closed and $\mathbb{F}$–invariant subset of $\partial^2\mathbb{F}$. Correspondingly, the **counting current** of a (nontrivial) conjugacy class $\alpha = \beta^m$, where β is not a proper power, is defined as

$$\eta_\alpha := m\eta_\beta := m \sum_{b \in \beta} \delta_{(b^-,b^+)} + \delta_{(b^+,b^-)}.$$

One may check that η_α is indeed a current and moreover that $\mathrm{supp}(\eta_\alpha) = L(\alpha)$.

2.3 Trees

An $\mathbb{R}$–**tree** is a 0–hyperbolic geodesic metric space. Alternately, an $\mathbb{R}$–tree is a metric space in which there is a unique embedded path between any two points and this path is a geodesic. Throughout this paper, we will use the term **tree** to mean an $\mathbb{R}$–tree equipped with an isometric and minimal action of $\mathbb{F}$; a tree is **minimal** if it does not contain a proper $\mathbb{F}$–invariant subtree.

We write $\ell_T(a)$ for the **translation length** of an element $a \in \mathbb{F}$ acting on a tree T, that is, $\ell_T(a) = \inf_{t \in T} d(t, a \cdot t)$. Notice that $\ell_T(a)$ depends only on the conjugacy class α of a. The element a acts **hyperbolically** (with an invariant axis) on T if $\ell_T(\alpha) > 0$ and **elliptically** (with a nonempty fixed subtree) if $\ell_T(\alpha) = 0$. The tree T is said to be **free** if $\ell_T(\alpha) > 0$ for all nontrivial conjugacy classes α.

Coulbois, Hilion and Lustig [CHL] have associated to every tree T a **dual lamination**

$$L(T) := \bigcap_{\varepsilon > 0} \left(\overline{\bigcup_{\alpha \in \Omega_\varepsilon(T)} L(\alpha)} \right) \subset \partial^2 \mathbb{F},$$

where $\Omega_\varepsilon(T) = \{\alpha : \ell_T(\alpha) < \varepsilon\}$ is the set of conjugacy classes with short translation length in T and the closure is taken in $\partial^2 \mathbb{F}$. Observe that $L(\alpha) \subset L(T)$ if and only if $\ell_T(\alpha) = 0$; thus T is free if and only if $L(T)$ does not contain the lamination $L(\alpha)$ of any nontrivial conjugacy class α. The set $L(T)$ is nonempty, and thus a bona fide lamination, unless T is free and simplicial. We refer the reader to [CHL] for a more detailed discussion of $L(T)$.

We say that a tree is **very small** if the stabilizer of any segment of T is maximal cyclic and the stabilizer of every tripod is trivial [CL]. A tree T is said to have **dense orbits** if every $\mathbb{F}$ orbit is dense in T. At the other extreme, if every orbit is discrete then the tree is said to be **simplicial**. A tree is **arational** if there does not exist a proper free factor A of $\mathbb{F}$ and an A–invariant subtree on which A acts with dense orbits. Following Guirardel [Gui], we say that a tree T is **indecomposable** if for every pair of nondegenerate arcs $\tau, \tau' \subset T$ there exist $a_1, \ldots, a_n \in \mathbb{F}$ so that $\tau' \subset a_1 \tau \cup \cdots \cup a_n \tau$ with $a_i \tau \cap a_{i+1} \tau$ nondegenerate for each $1 \leq i < n$; indecomposablity is thus a strong mixing property for the action of $\mathbb{F}$ on T. The following theorem of Reynolds clarifies the relationship between these notions:

Theorem 2.2 (Reynolds [Rey2])**.** *A minimal, very small tree is arational if and only if it is indecomposable and either* (1) *free or* (2) *dual to a filling measured lamination on a once-punctured surface.*

2.4 The free factor complex

A nontrivial subgroup $A \leq \mathbb{F}$ of $\mathbb{F}$ is a **free factor** of $\mathbb{F}$ if there exists a complementary nontrivial subgroup $B \leq \mathbb{F}$ such that $\mathbb{F} = A * B$. As is common we often blur the distinction between free factors and their conjugacy classes. The **(free) factor complex** of $\mathbb{F}$ is the simplicial complex $\mathcal{F}$ whose k–simplices consist of chains $A_0 < \cdots < A_k$ of properly nested (conjugacy classes of) free factors of $\mathbb{F}$ and whose face inclusions correspond to subchains. Note that $\mathcal{F}$ is not locally compact, and that the group $\mathrm{Out}(\mathbb{F})$ acts on $\mathcal{F}$ by simplicial automorphisms. We equip $\mathcal{F}$ with the path metric in which simplices are all isometric to standard Euclidean simplices with side lengths equal to 1; the induced path metric on the 1–skeleton makes $\mathcal{F}^1$ into a simplicial metric graph with all edges having length 1. With this setup we have the following foundational result of Bestvina and Feighn:

Theorem 2.3 (Bestvina–Feighn [BF2])**.** *The factor complex $\mathcal{F}$ is Gromov hyperbolic.*

As the full complex $\mathcal{F}$ and simplicial graph $\mathcal{F}^1$ are quasi-isometric, we henceforth work exclusively with the 1–skeleton. In particular, for $A, B \in \mathcal{F}^0$ we write $d_{\mathcal{F}}(A, B)$ to mean the distance from A to B in the path space $\mathcal{F}^1$.

2.5 Outer space

Let $\mathcal{R}$ be the $\mathrm{rk}(\mathbb{F})$–petal rose with base vertex v, and fix an isomorphism $\pi_1(\mathcal{R}, v) \cong \mathbb{F}$. A **core graph** is a finite 1–dimensional CW-complex G with no valence 1 vertices; and by a **metric** on G we mean a path metric for which the interior of each 1–cell (with the induced path metric) is isometric to a positive-length open subinterval of $\mathbb{R}$. The **volume** of G is the sum of its edge lengths, and a **marking** of G is a homotopy equivalence $g \colon \mathcal{R} \to G$. Culler and Vogtmann's [CV] **unprojectivized Outer space** of $\mathbb{F}$ is the space $cv = \{(G, g)\}$ of marked metric core graphs, modulo the equivalence relation $(G, g) \sim (H, h)$ if there exists an isometric **change of marking map** $G \to H$ in the homotopy class $h \circ g^{-1}$. By (projectivized) **Outer space** $\mathcal{X}$, we simply mean the subset of cv consisting of volume 1 marked metric graphs. We equip $\mathcal{X}$ with the **asymmetric metric** $d_{\mathcal{X}}$ defined as follows:

$$d_{\mathcal{X}}((G, g), (H, h)) := \inf \left\{ \log(\mathrm{Lip}(\phi)) \mid \phi \simeq h \circ g^{-1} \right\}$$

where $\mathrm{Lip}(\phi)$ denotes the optimal Lipschitz constant for the change of marking map $\phi \colon G \to H$. The **symmetrization** $d_{\mathcal{X}}^{\mathrm{sym}}(G, H) = d_{\mathcal{X}}(G, H) + d_{\mathcal{X}}(H, G)$ is an honest metric and defines the topology on $\mathcal{X}$. We will suppress the marking and metric and denote points in cv and $\mathcal{X}$ simply by the underlying graph.

Given any subgroup $A \leq \mathbb{F}$ (or conjugacy class thereof) and point $G \in cv$, we write $A|G$ for the maximal core subgraph of the cover of G corresponding to A and we equip this with the pull-back metric from the immersion $A|G \to G$. When convenient we will blur the distinction between the metric core graph $A|G$ and the immersion $A|G \to G$ itself. For a conjugacy class α and write $\ell(\alpha|G)$ for the volume of the graph $\alpha|G$; this is the **length of α at G**. With this notation we have the following useful formula for the metric [FM]:

$$d_{\mathcal{X}}(G, H) = \log \left(\sup_{\alpha \in \mathbb{F}} \frac{\ell(\alpha|H)}{\ell(\alpha|G)} \right).$$

Observe that the universal cover $\tilde{G}$ of a graph $G \in cv$ is naturally a simplicial $\mathbb{R}$–tree equipped with an action of $\mathbb{F} \cong \pi_1(G)$ by deck transformations (where the isomorphism $\mathbb{F} \cong \pi_1(G)$ is provided by the marking $\mathcal{R} \to G$). With this perspective $\ell(\alpha|G)$ is simply the translation length $\ell_{\tilde{G}}(\alpha)$ of the conjugacy class α on the tree $\tilde{G}$. In fact, this correspondence gives a bijection between cv and the set of free simplicial $\mathbb{R}$–trees up to $\mathbb{F}$–equivariant isometry.

Asymmetry in Outer space. Care must be taken to cope with the asymmetry inherent in Outer space. For us a **geodesic in** $\mathcal{X}$ always means a *directed* geodesic, that is, a map $\gamma \colon \mathbf{I} \to \mathcal{X}$ so that $d_{\mathcal{X}}(\gamma(s), \gamma(t)) = t - s$ for all $s < t$. Similarly a K–**quasigeodesic** is a map $\gamma \colon \mathbf{I} \to \mathcal{X}$ so that

$$\tfrac{1}{K}(t - s) - K \leq d_{\mathcal{X}}(\gamma(s), \gamma(t)) \leq K(t - s) + K$$

for all $s < t$. On the other hand, for $r > 0$ the r–**neighborhood** $\mathcal{N}_r(U)$ of a subset $U \subset \mathcal{X}$ must be defined using the symmetrized metric:

$$\mathcal{N}_r(U) := \{ G \in \mathcal{X} \mid d_{\mathcal{X}}^{\mathrm{sym}}(G, u) \leq r \text{ for some } u \in U \}.$$

The **Hausdorff distance** $d_{\mathrm{Haus}}(U, W)$ between two subsets $U, W \subset \mathcal{X}$ is then defined, as usual, to be the infimal r so that $U \subset \mathcal{N}_r(W)$ and $W \subset \mathcal{N}_r(U)$. For $\varepsilon > 0$, we write

$$\mathcal{X}_\varepsilon := \{ G \in \mathcal{X} \mid \ell(\alpha|G) \geq \varepsilon \text{ for all } \alpha \in \mathbb{F} \setminus 1 \}$$

for the ε–**thick part** of Outer space. The following important result bounds the asymmetry in $\mathcal{X}_\varepsilon$.

Lemma 2.4 (Handel–Mosher [HM], Algom-Kfir–Bestvina [AKB]). *For every $\varepsilon > 0$ there exists $\mathsf{M}_\varepsilon \geq 1$ so that for all $G, H \in \mathcal{X}_\varepsilon$ one has*

$$d_{\mathcal{X}}(H, G) \leq d_{\mathcal{X}}^{\mathrm{sym}}(H, G) = d_{\mathcal{X}}^{\mathrm{sym}}(G, H) \leq \mathsf{M}_\varepsilon \cdot d_{\mathcal{X}}(G, H).$$

Projecting to the factor complex. There is a coarse projection $\pi_{\mathcal{F}}\colon \mathcal{X} \to \mathcal{F}$ defined by sending $G \in \mathcal{X}$ to the set of free factors $\pi_1(G')$ corresponding to proper, connected, noncontractible subgraphs G' of G (here $\pi_1(G')$ is identified with a free factor of $\mathbb{F}$ by the marking $\mathcal{R} \to G$). One may easily check that $\mathrm{diam}_{\mathcal{F}}(\pi_{\mathcal{F}}(G)) \leq 4$ [BF2, Lemma 3.1], so we are justified in viewing $\pi_{\mathcal{F}}$ as a coarse projection. For $G, H \in \mathcal{X}$ we define

$$d_{\mathcal{F}}(G,H) = \mathrm{diam}_{\mathcal{F}}(\pi_{\mathcal{F}}(G) \cup \pi_{\mathcal{F}}(H)).$$

The following appears as Lemma 2.9 in [DT] and follows from [BF2, Corollary 3.5].

Lemma 2.5. *For all $G, H \in \mathcal{X}$ we have $d_{\mathcal{F}}(G,H) \leq 80 d_{\mathcal{X}}(G,H) + 80$.*

The projection $\pi_{\mathcal{F}}$ provides an important connection between the geometries of $\mathcal{X}$ and $\mathcal{F}$. For example, the following stability result uses the geometry of $\mathcal{F}$ to establish aspects of hyperbolicity in $\mathcal{X}$ and served as a main tool in our proof of Theorem 1.1.

Theorem 2.6 (Dowdall–Taylor [DT]). *Let $\gamma\colon \mathbf{I} \to \mathcal{X}$ be a K–quasigeodesic whose projection $\pi \circ \gamma\colon \mathbf{I} \to \mathcal{F}$ is also a K–quasigeodesic. Then there exist constants $A, \varepsilon > 0$ and $K' \geq 1$ depending only on K (and the injectivity radius of the terminal endpoint $\gamma(\mathbf{I}_+)$ when $\mathbf{I}_+ < \infty$) with the following property: If $\rho\colon \mathbf{J} \to \mathcal{X}$ is any geodesic with the same endpoints as γ, then*

(i) $\gamma(\mathbf{I}), \rho(\mathbf{J}) \subset \mathcal{X}_{\varepsilon}$,

(i) $d_{\mathrm{Haus}}(\gamma(\mathbf{I}), \rho(\mathbf{J})) < A$, and

(ii) $\pi \circ \rho\colon \mathbf{J} \to \mathcal{F}$ is a (parameterized) K'–quasigeodesic.

Folding. We will need a particular class of directed geodesics in $\mathcal{X}$ called folding paths, which we now briefly describe. A **segment** in a metric core graph G is a locally isometric immersion of an interval $[0, L]$ into G, and a **direction** at $p \in G$ is a germ of nondegenerate segments with $0 \mapsto p$. A **turn** is an unordered pair $\{d, d'\}$ of distinct directions at a vertex of G.

A map $\phi\colon G \to H$ of metric core graphs that is a local $\mathrm{Lip}(\phi)$–homothety induces a derivative map D_{ϕ} which sends a direction at p to a direction at $\phi(p)$. Two directions at p are said to be in the same **gate** if they are identified by D_{ϕ}. The map then gives rise to an **illegal turn structure** on G, whereby a turn $\{d, d'\}$ is **illegal** if d and d' are in the same gate and is **legal** otherwise.

We say that a map $\phi\colon G \to H$ between points $G, H \in \mathcal{X}$ is a **folding map** if it is homotopic to the change of markings $h \circ g^{-1}$, is a local $\mathrm{Lip}(\phi)$–homothety, satisfies $d_{\mathcal{X}}(G,H) = \log \mathrm{Lip}(\phi)$, and it induces at least 2 gates at each point $p \in G$. As described in [BF2, §2], each folding map $\phi\colon G \to H$ gives rise to a unique **folding path** $\{G_t\}_{t \in [0,L]}$ in $\mathcal{X}$ via first *folding all illegal turns at speed one* and second *rescaling* to obtain graphs in $\mathcal{X}$. The folding path comes equipped with a family of folding maps $\{\phi_{st}\colon G_s \to G_t\}_{s \leq t}$ satisfying

$$\phi_{0L} = \phi, \quad \phi_{ss} = \mathrm{Id}_{G_s}, \quad \text{and} \quad \phi_{rt} = \phi_{st} \circ \phi_{rs} \text{ with } \phi_{rs} \text{ and } \phi_{rt} \text{ inducing the same illegal turn structure on } G_r$$

for all $0 \leq r \leq s \leq t \leq L$. Our convention is to parameterize the folding path $\{G_t\}_{t \in [0,L]}$ with respect to arc length in Outer space. With this convention, the assignment $t \mapsto \gamma(t) = G_t$ defines a directed geodesic $\gamma\colon [0,L] \to \mathcal{X}$ from $G = G_0$ to $H = G_L$. See Proposition 2.2 and Notation 2.4 of [BF2] or [DT, §2] for the details of this construction; further properties will be recalled in §6.2.

2.6 The boundaries of Outer space and the factor complex

The length functions give an embedding of unprojectivized Outer space cv into $\mathbb{R}^{\mathbb{F}}$ via $G \mapsto (\ell(a|G))_{a \in \mathbb{F}}$, and thus cv inherits the subspace topology from $\mathbb{R}^{\mathbb{F}}$. The resulting topology on $\mathcal{X} \subset cv$ agrees with the one induced by the symmetrized metric $d_{\mathcal{X}}^{\mathrm{sym}}$. The work of Cohen–Lustig [CL] and Bestvina–Feighn [BF1] shows that the closure $\overline{cv}$ of cv in $\mathbb{R}^{\mathbb{F}}$ may be identified with the space of minimal very small trees. Projectivizing, one similarly identifies the closure $\overline{\mathcal{X}}$ of $\mathcal{X} \subset \mathbb{PR}^{\mathbb{F}}$ with the space of projective classes of minimal very small

trees. The **boundary of** $\mathfrak{X}$ is consequently defined to be the set $\partial\mathfrak{X} := \overline{\mathfrak{X}} \setminus \mathfrak{X}$ of projective classes very small trees that are not both free and simplicial.

In [KL1], Kapovich and Lustig introduced an Out($\mathbb{F}$)–invariant intersection pairing $\langle\cdot,\cdot\rangle$ between very small minimal trees and currents; we record here a few of its properties:

Theorem 2.7 (Kapovich–Lustig [KL1, KL2])**.** *There is a unique* Out($\mathbb{F}$)–*invariant continuous pairing*

$$\langle\cdot,\cdot\rangle\colon \overline{cv} \times \mathrm{Curr}(\mathbb{F}) \to \mathbb{R}_+$$

which is homogeneous in the first coordinate and linear in the second. Moreover, for every tree T, current μ, and conjugacy class α, we have that $\langle T,\eta_\alpha\rangle = \ell_T(\alpha)$ and that $\langle T,\mu\rangle = 0$ if and only if $\mathrm{supp}(\mu) \subset L(T)$.

We will be particularly interested in the case where the tree T is free and indecomposable.

Theorem 2.8 (Coulbois–Hilion–Reynolds [CHR, Corollary 1.4])**.** *If $T \in \overline{cv}$ is free and indecomposable, then $\langle T,\mu\rangle = 0$ if and only if $\mathrm{supp}(\mu) = L'(T)$.*

Finally, let $\mathcal{AT}$ be the subspace of ∂cv consisting of arational trees. For $T,'T \in \mathcal{AT}$, say $T \sim T'$ if $L(T) = L(T')$. The following result computes the boundary of $\mathcal{F}$.

Theorem 2.9 (Bestvina–Reynolds [BR], Hamenstädt [Ham3])**.** *The map $\pi_\mathcal{F}\colon \mathfrak{X} \to \mathcal{F}$ has a continuous extension to a map $\partial\pi_\mathcal{F}\colon \mathcal{AT} \to \partial\mathcal{F}$, in the sense that if $G_i \to T$ in $\overline{\mathfrak{X}}$ and $T \in \mathcal{AT}$, then $\pi_\mathcal{F}(G_i) \to \partial\pi_\mathcal{F}(T)$ in $\mathcal{F} \cup \partial\mathcal{F}$. Moreover, if $T \sim T'$ then $\partial\pi_\mathcal{F}(T) = \partial\pi_\mathcal{F}(T')$, and the induced map $\mathcal{AT}/\sim\,\to \partial\mathcal{F}$ is a homeomorphism.*

3 Alignment preserving maps and boundaries

In §4 we will introduce and analyze a new Out($\mathbb{F}$)–graph termed the co-surface graph. For this analysis, we develop a general framework for computing the boundary of any space whose hyperbolicity may be obtained by the Kapovich–Rafi method [KR]. As this result is applicable in other contexts, we state it in general terms.

Let X and Y be geodesic metric spaces. We say that three (ordered) points $a,b,c \in X$ are K**–aligned** if $d_X(a,b) + d_X(b,c) \le d_X(a,c) + K$; the points are simply said to be **aligned** if they are 0–aligned. We say that a Lipschitz map $p\colon X \to Y$ is **alignment preserving** if there exists $K \ge 0$ such that $p(a),p(b),p(c)$ are K–aligned whenever a,b,c are aligned.

Lemma 3.1. *Suppose that $q\colon W \to X$ and $p\colon X \to Y$ are alignment preserving maps between geodesic metric spaces and that X is δ–hyperbolic. Then for all $L \ge 0$ there is an $L' \ge 0$ such that $p(a),p(b),p(c)$ are L'–aligned whenever a,b,c are L–aligned. Moreover, the composition $p \circ q\colon W \to Y$ is alignment preserving.*

Let us formalize some observations that will aid in the proof. Firstly, every δ–thin triangle admits a 2δ**–barycenter**, meaning a point ω that lies within 2δ of each side of the triangle (and is consequently 4δ–aligned between any two vertices). Secondly, whenever the three triples (a,ω,b), (b,ω,c), and (c,ω,a) are each K–aligned, the triangle inequality immediately gives

$$|d(a,\omega) - (b|c)_a| \le K. \tag{2}$$

Combining these, we see that in a δ–hyperbolic space X, the Gromov product $(b|c)_a$ lies within 4δ of $d_X(a,\omega)$ for any 2δ–barycenter ω of the geodesic triangle $\triangle(a,b,c)$.

Proof. Let $K \ge 0$ be the alignment constant of the alignment preserving map p. Take L–aligned points a,b,c in X and let ω be a 2δ–barycenter for the geodesic triangle $\triangle(a,b,c)$. Then $d_X(b,\omega)$ is within 4δ of $(a|c)_b$, which is in turn bounded by $L/2$ since a,b,c are L–aligned. Thus there is a point $x \in [a,c]$ with $d_X(b,x) \le 6\delta + L$. Then $p(a),p(x),p(c)$ are K–aligned, and so the points $p(a),p(b),p(c)$ are L'–aligned, where L' is K plus $2(6\delta + L)$ times the Lipschitz constant for p. This proves the first claim. The second claim now follows immediately from the first. $\square$

If X and Y are hyperbolic and $p\colon X \to Y$ is alignment preserving, the Y–**subboundary of** X (relative to p) is defined to be

$$\partial_Y X = \{\gamma(\infty) \in \partial X \mid \gamma\colon \mathbb{R}_+ \to X \text{ is a quasigeodesic ray with } \operatorname{diam}_Y p(\gamma(\mathbb{R}_+)) = \infty\}.$$

Informally, $\partial_Y X$ consists of those points in ∂X that "project to infinity" in Y. This is made precise by the following theorem.

Theorem 3.2 (Boundaries). *Suppose that $p\colon X \to Y$ is a coarsely surjective, alignment preserving map between hyperbolic spaces. Then p admits an extension to a homeomorphism $\partial p\colon \partial_Y X \to \partial Y$. Moreover, the extension $p \cup \partial p\colon X \cup \partial_Y X \to Y \cup \partial Y$ is continuous in the sense that if $x_n \to \lambda \in \partial_Y X$ as $n \to \infty$, then $p(x_n) \to \partial p(\lambda) \in \partial Y$.*

Proof. Let δ be the hyperbolicity constant of X, let L be the Lipschitz constant for p, and let $K' \geq 0$ be the constant, provided by Lemma 3.1, such that $p(a), p(b), p(c)$ are K'–aligned whenever $a, b, c \in X$ are 4δ–aligned. We then have the following useful observation: If ω is any 2δ–barycenter for an arbitrary triangle $\triangle(a_1, a_2, a_3)$ in X, then the triples $(p(a_i), p(\omega), p(a_j))$ for distinct $i, j \in \{1, 2, 3\}$ are each K'–aligned and so we may apply (2) in both X and Y to conclude that

$$(p(a_2)|p(a_3))_{p(a_1)} \leq K' + d_Y(p(a_1), p(\omega)) \leq K' + L d_X(a_1, \omega) \leq L(a_2|a_3)_{a_1} + 4\delta + K'. \tag{3}$$

To define the map ∂p, choose a quasigeodesic ray $\gamma\colon \mathbb{R}_+ \to X$ with $\operatorname{diam}_Y p(\gamma(\mathbb{R}_+)) = \infty$ and consider the admissible sequence $\{a_n\}_{n=0}^{\infty}$, where $a_n = \gamma(n)$. Set $b_n = p(a_n)$. Since ordered triples of points along $\gamma(\mathbb{R}_+)$ are uniformly aligned by Proposition 2.1, the assumption $\operatorname{diam}_Y p(\gamma(\mathbb{R}+)) = \infty$ in fact implies that $\lim_t d_Y(b_0, p(\gamma(t))) = \infty$. For each pair $n, m \geq 0$, choose a 2δ–barycenter $c_{n,m}$ for the triangle $\triangle(a_0, a_n, a_m)$. Then $(a_n|a_m)_{a_0}$ is within 4δ of $d_X(a_0, c_{n,m})$ by (2). By Proposition 2.1, $c_{n,m}$ also lies within uniformly bounded distance of $\gamma(t_{n,m})$ for some $t_{n,m} \in \mathbb{R}_+$. By admissibility and the fact that γ is a quasigeodesic, the quantities $d_X(a_0, c_{n,m})$ and $t_{n,m}$ both tend to infinity as $n, m \to \infty$. Therefore $\lim_{n,m} d_Y(b_0, p(c_{n,m})) = \infty$ since p is Lipschitz. However, $(b_n|b_m)_{b_0}$ is within K' of $d_Y(b_0, p(c_{n,m}))$ because $p(c_{n,m})$ is K'–aligned between the three points b_0, b_n and b_m. Consequently $\{b_n\}$ is admissible, and we may define $\partial p(\gamma(\infty))$ to be $\{b_n\} \in \partial Y$.

We now prove that ∂p is well-defined and that $p \cup \partial p$ is continuous. Let $\lambda = \gamma(\infty) \in \partial_Y X$ with γ, $\{a_n\}$, and $\{b_n\}$ as above. Suppose that $\{x_n\}$ is a sequence in X converging to λ. This simply means that $\{x_n\}$ is admissible and equivalent to $\{a_n\}$. Letting $e_{n,m}$ denote a 2δ–barycenter for $\triangle(a_0, a_n, x_m)$, we have that $d_X(a_0, e_{n,m})$ is within 4δ of $(a_n|x_m)_{a_0}$ and thus tends to infinity. This barycenter $e_{n,m}$ is also uniformly close, again by Proposition 2.1, to some point $\gamma(s_{n,m})$ with $s_{n,m}$ necessarily tending to infinity since γ is a quasigeodesic. As before, it follows that $d_Y(b_0, p(e_{n,m})) \to \infty$ and, since p is alignment preserving, that this quantity coarsely agrees with $(p(a_n)|p(x_m))_{b_0}$. Thus $\{p(x_n)\}$ is equivalent to $\{p(a_n)\}$, proving that $\{p(x_n)\}$ is admissible and converges to $\partial p(\lambda) = \{b_n\}$. In particular, for any quasigeodesic $\gamma'\colon \mathbb{R}_+ \to X$ with $\gamma'(\infty) = \lambda$, it follows that $\{p(\gamma'(n))\}$ is equivalent to $\{b_n\} = \{p(\gamma(n))\}$. Thus ∂p is well-defined and the extension $p \cup \partial p$ is continuous in the manner claimed.

We next show ∂p is injective. Suppose $\lambda, \mu \in \partial_Y X$ satisfy $\partial p(\lambda) = \partial p(\mu)$. If $x_n, z_n \in X$ are any sequences with $x_n \to \lambda$ and $z_n \to \mu$, then $\{p(x_n)\}$ and $\{p(z_n)\}$ are equivalent by the continuity of $p \cup \partial p$. Therefore $(p(x_n)|p(z_m))_{p(x_0)} \to \infty$ which, by (3), forces $(x_n|z_m)_{x_0} \to \infty$ as well. Thus $\lambda = \mu$ and ∂p is injective.

To see surjectivity, let $\eta \in \partial Y$. For a sequence $y_n \in Y$ with $y_n \to \eta$, choose points $x_n \in X$ so that $d_Y(p(x_n), y_n)$ is uniformly bounded and thus $p(x_n) \to \eta$ as well. Since $\{p(x_n)\}$ is admissible (3) implies that $\{x_n\}$ is admissible and so converges to some point $\lambda \in \partial X$. Now build 10δ–quasigeodesic $\gamma\colon \mathbb{R}_+ \to X$ with $\gamma(0) = x_0$ and $\gamma(\infty) = \lambda$. Refer to Figure 1. Also let $e_{n,m}$ be a 2δ–barycenter of $\triangle(x_0, x_n, x_m)$. We claim that the projection of γ to Y has infinite diameter. To see this, fix any $D > 0$ and use admissibility of $\{p(x_n)\}$ to find N so that $d_Y(p(x_0), p(e_{n,m})) > D$ for all $n, m \geq N$. Fix some $k \geq N$. Using the equivalence of $\{x_n\}$ and $\{\gamma(n)\}$, we may then choose $m \geq N$ so that $d_X(x_0, c_m) \geq d_X(x_0, x_k) + 8\delta$, where c_m denotes a 2δ–barycenter for $\triangle(x_0, x_m, \gamma(m))$. Since the triple $(x_0, e_{k,m}, x_k)$ is 4δ–aligned, this implies $d_X(x_0, c_m) \geq d_X(x_0, e_{k,m}) + 4\delta$. Choosing points e' and c' in $[x_0, x_m]$ within 2δ of $e_{k,m}$ and c_m, respectively, it follows that $d_X(x_0, c') \geq d_X(x_0, e')$ so that $e' \in [x_0, c']$. Now, since c_m is a barycenter for $\triangle(x_0, x_m, \gamma(m))$, there

9
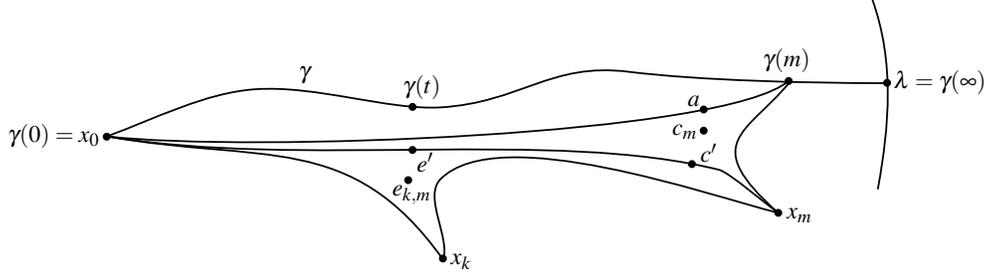
Figure 1: Proving surjectivity of ∂p.

is a point $a \in [x_0, \gamma(m)]$ with $d_X(c', a) \leq 4\delta$. By thinness of the triangle $\triangle(x_0, c', a)$, we see that $e' \in [x_0, c']$ lies within 5δ of $[x_0, a] \subset [x_0, \gamma(m)]$. Therefore, since $d_X(e_{k,m}, e') \leq 2\delta$ and γ is a quasigeodesic, Proposition 2.1 and the triangle inequality imply that $e_{k,m}$ is within $7\delta + R_0(10\delta, \delta)$ of some point $\gamma(t)$. Consequently, $d_Y(p(x_0), p(\gamma(t)))$ coarsely agrees with $d_Y(p(x_0), p(e_{k,m})) \geq D$ up to uniformly bounded additive error. Since D here is arbitrary, it follows that $\mathrm{diam}_Y(p(\gamma)) = \infty$. Thus $\lambda \in \partial_Y X$, which proves that ∂p is surjective.

Finally we prove ∂p is a homeomorphism. Firstly, for any $\lambda, \mu \in \partial_Y X$ and $x \in X$, Equation (3) and our proof of surjectivity show that $(\partial p(\lambda) | \partial p(\mu))_{p(x)} \leq L(\lambda | \mu)_x + 4\delta + K'$. By the definition of the topology on the Gromov boundary, it immediately follows that $\partial p \colon \partial_Y X \to \partial Y$ is open. Conversely, for any $\eta \in \partial Y$ and $D > 0$, we may, as above, build a 10δ–quasigeodesic $\gamma \colon \mathbb{R}_+ \to X$ converging to $\lambda = (\partial p)^{-1}(\eta)$ and whose projection to Y has infinite diameter. Set $x = \gamma(0)$. Thus there is some R so that $d_Y(p(x), p(z)) \geq D + K'$ for all $z \in X$ that lie within 2δ of γ and satisfy $d_X(x, z) \geq R$. Now if $\mu \in \partial_Y X$ is such that $(\mu | \lambda)_x \geq R + 4\delta$, then we may choose a sequence $\{z_n\}$ converging to μ so that

$$\liminf_{n,m \to \infty} (z_n | \gamma(m))_x \geq R + 4\delta.$$

Thus if $e_{n,m}$ is a 2δ–barycenter for $\triangle(x, z_n, \gamma(m))$, then $\liminf_{n,m} d_X(x, e_{n,m}) \geq R$. But since $e_{n,m}$ lies within 2δ of γ, we have that $d_Y(p(x), p(e_{n,m})) \geq D + K'$ and thus also $(p(z_n) | p(\gamma(m)))_{p(x)} \geq D$ for all large n, m. This proves that $(\partial p(\mu) | \eta)_{p(x)} \geq D$ for all μ satisfying $(\mu | \lambda)_x \geq R + 4\delta$. Therefore ∂p is continuous. $\square$

Remark 3.3. Note that if the hypothesis of coarse surjectivity in Theorem 3.2 is dropped, the proof shows that the map $\partial p \colon \partial_Y X \to \partial Y$ is a topological embedding.

We also record the following useful lemma, the idea of which is well-known to experts (see for example [Ham2, Lemma 2.6]). First say that $p \colon X \to Y$ is **metrically proper** if for any $D \geq 0$ there is a $C \geq 0$ so that $d_X(a, b) \geq C$ implies $d_Y(p(a), p(b)) \geq D$.

Lemma 3.4. *Suppose that X and Y are geodesic metric spaces. If $p \colon X \to Y$ is alignment preserving and metrically proper, then p is a quasi-isometric embedding.*

Proof. Since $p \colon X \to Y$ is alignment preserving, there is a constant K such that whenever $a, b, c \in X$ are 0–aligned we have

$$d_Y(p(a), p(c)) \geq d_Y(p(a), p(b)) + d_Y(p(b), p(c)) - K.$$

Further, since $p \colon X \to Y$ is metrically proper, there is a $C > 0$ such that if $a, b \in X$ satisfy $d_X(a, b) \geq C$, then $d_Y(p(a), p(b)) \geq 2K$. Now let x and x' be points of X with $d_X(x, x') = d$ and let $\gamma \colon [0, d] \to X$ be a geodesic

with $\gamma(0) = x$ and $\gamma(d) = x'$. Let N be the largest integer less than $\frac{d}{C}$, and set $a_i = \gamma(iC)$ for $0 \le i \le N$. Then,

$$d_Y(p(x), p(x')) \ge d_Y(p(a_0), p(a_N)) + d_Y(p(a_N), p(x')) - K$$

$$\ge d_Y(p(a_N), x') - K + \sum_{i=0}^{N-1}(d_Y(p(a_i), p(a_{i+1})) - K)$$

$$\ge -K + K \cdot N$$

$$\ge \frac{K}{C} \cdot d_X(x, x') - 2K.$$

Since $p \colon X \to Y$ is Lipschitz by assumption, this completes the proof. $\qquad\square$

4 The co-surface graph $\mathcal{CS}$

This section introduces the co-surface graph $\mathcal{CS}$ of the free group $\mathbb{F}$ and develops its basic properties. First we define $\mathcal{CS}$ and discuss its relationship to other $\mathrm{Out}(\mathbb{F})$–graphs appearing in the literature. Then in §4.1 we use the theory of alignment preserving maps to calculate the boundary of $\mathcal{CS}$. Finally, in §4.2 we show that a subgroup $\Gamma \le \mathrm{Out}(\mathbb{F})$ qi-embeds into $\mathcal{CS}$ if and only if it is purely atoroidal and qi-embeds into $\mathcal{F}$.

Recall that an element $a \in \mathbb{F}$ is **primitive** is a belongs to a free basis for $\mathbb{F}$; this is an invariant of the conjugacy class α of a and so we also call α primitive. The **primitive loop graph** of $\mathbb{F}$ is the simplicial graph $\mathcal{PL}$ whose vertices are the primitive conjugacy classes of $\mathbb{F}$, and where two conjugacy classes α, β are joined by an edge if and only if they have representatives that are jointly part of a free basis of $\mathbb{F}$. We equip $\mathcal{PL}$ with the path metric $d_{\mathcal{PL}}$ in which each edge has length 1. As each primitive element generates a cyclic free factor of $\mathbb{F}$, there is natural inclusion map $\mathcal{PL}^0 \to \mathcal{F}$. It is straightforward to check that this inclusion is 2–bilipschitz and 1–dense and therefore admits a 4–quasi-isometry coarse inverse which we denote $\mathcal{D} \colon \mathcal{F} \to \mathcal{PL}$. We also have the coarse projection $\pi_{\mathcal{PL}} \colon \mathcal{X} \to \mathcal{PL}$ defined by sending $G \in \mathcal{X}$ to the set of embedded closed loops on G; this projection coarsely agrees with the composition $\mathcal{D} \circ \pi_{\mathcal{F}}$.

The primitive loop graph measures, in a sense, how algebraically complicated primitive conjugacy classes are with respect to each other. The co-surface graph, on the other hand, is designed to measure how *topologically* complicated primitive conjugacy classes are with respect to each other:

Definition 4.1 (Co-surface graph $\mathcal{CS}$). The **co-surface graph** $\mathcal{CS}$ of the free group $\mathbb{F}$ is the simplicial graph whose vertices are conjugacy classes of primitive elements, and where two vertices α and β are joined by an edge if there is a once-punctured surface S and an isomorphism $\pi_1(S) \cong \mathbb{F}$ with respect to which α and β may both be represented by *simple* closed curves on S.
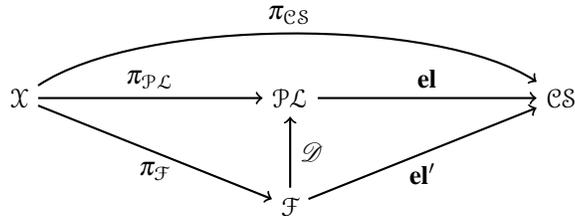
Figure 2: The co-surface graph and its friends. All rightward pointing arrows are coarsely Lipschitz, and $\mathcal{D}$ is a quasi-isometry. Coarsely speaking, all maps are $\mathrm{Out}(\mathbb{F})$–equivariant and the diagram commutes.

In other words, each once-punctured surface S with $\pi_1(S) \cong \mathbb{F}$ determines a subset $C_S \subset \mathcal{PL}^0$ consisting of those primitive conjugacy classes that correspond to (nonseparating) simple closed curves on S. As our graph $\mathcal{CS}$ is obtained by collapsing each $C_S \subset \mathcal{PL}$ to a set of diameter 1, it records the geometry of the

primitive conjugacy classes that remains after all the "surface sets" C_S have been crushed—hence the name "co-surface" graph. We equip $\mathcal{CS}$ with the path metric $d_{\mathcal{CS}}$ in which each edge has length 1 and note that $\mathrm{Out}(\mathbb{F})$ acts simplicially (and hence isometrically) on $\mathcal{CS}$.

From basic topology, we know that if primitive elements $a, b \in \mathbb{F}$ are jointly part of a free basis of $\mathbb{F}$, then one may build a once-punctured surface S and an isomorphism $\pi_1(S) \cong \mathbb{F}$ under which a and b correspond to disjoint simple closed curves on S. Therefore the "identity" map $\mathcal{PL}^0 \to \mathcal{CS}^0$ extends to a simplicial and hence 1–Lipschitz $\mathrm{Out}(\mathbb{F})$–equivariant "electrification" map $\mathbf{el} \colon \mathcal{PL} \to \mathcal{CS}$. Define $\mathbf{el}' \colon \mathcal{F} \to \mathcal{CS}$ to be the composition $\mathbf{el}' = \mathbf{el} \circ \mathscr{D}$. The purpose of this section is to establish the following essential properties of $\mathcal{CS}$. For the statement, recall that $\phi \in \mathrm{Out}(\mathbb{F})$ is **fully irreducible** if no positive power of ϕ fixes any conjugacy class of proper free factors of $\mathbb{F}$.

Theorem 4.2 (Properties of $\mathcal{CS}$). *For the free group $\mathbb{F}$ of rank at least 3, the co-surface graph $\mathcal{CS}$ is hyperbolic and the map $\mathbf{el} \colon \mathcal{PL} \to \mathcal{CS}$ (and thus also $\mathbf{el}' \colon \mathcal{F} \to \mathcal{CS}$) is Lipschitz and alignment preserving. Moreover, $\phi \in \mathrm{Out}(\mathbb{F})$ acts as a loxodromic isometry of $\mathcal{CS}$ if and only if ϕ is atoroidal and fully irreducible.*

To establish these properties, we show that $\mathcal{CS}$ is quasi-isometric to another $\mathrm{Out}(\mathbb{F})$–graph that has appeared in several different forms in the literature under the name "intersection graph" (see Remark 4.5). To define this, say that a conjugacy class of $\mathbb{F}$ is **geometric** if it is either primitive or it corresponds to the cusp of a once-punctured surface whose fundamental group is identified with $\mathbb{F}$. Define the **intersection graph** to be the bipartite graph $\mathscr{I}$ whose vertices are geometric conjugacy classes and very small, simplicial, nonfree trees, and where a conjugacy class α is joined by an edge to a tree T if and only if $\ell_T(\alpha) = 0$. Note that there is an obvious action $\mathrm{Out}(\mathbb{F}) \curvearrowright \mathscr{I}$ and that the inclusion $\mathcal{PL}^0 \to \mathscr{I}^0$ extends to an $\mathrm{Out}(\mathbb{F})$–equivariant map $\mathcal{PL} \to \mathscr{I}$. Brian Mann and Patrick Reynolds have proven the following:

Theorem 4.3 (Mann [Man], Mann–Reynolds [MR1]). *The intersection graph is hyperbolic and the natural map $\mathcal{PL} \to \mathscr{I}$ is Lipschitz and alignment preserving.*

The proof of these results can be found in Mann's thesis [Man], which is publicly available through ProQuest. The main point of his proof is to show that the map $\mathcal{PL} \to \mathscr{I}$ fits the parameters of a recent theorem of Kapovich and Rafi [KR, Proposition 2.5]. Applying the Kapovich–Rafi result shows that hyperbolicity of the primitive loop graph implies hyperbolicity the intersection graph and, moreover, that the map $\mathcal{PL} \to \mathscr{I}$ is alignment preserving. One could use the same argument from [Man] to directly show that the Kapovich–Rafi result applies to the map $\mathbf{el} \colon \mathcal{PL} \to \mathcal{CS}$; rather than carry out this argument in detail, we simply invoke Theorem 4.3 and the following quasi-isometry between the co-surface graph and the intersection graph:

Proposition 4.4. *The graphs $\mathscr{I}$ and $\mathcal{CS}$ are $\mathrm{Out}(\mathbb{F})$–equivariantly quasi-isometric.*

Proof. In [Man, §2.4], Mann defined a graph $\mathscr{P}$ which he shows is quasi-isometric to $\mathscr{I}$. The vertices of $\mathscr{P}$ are marked $\mathrm{rk}(\mathbb{F})$–petal roses, and roses R and R' are joined by an edge if they have either a common petal or a common "cusp" (meaning that the given isomorphism $\pi_1(R) \cong \pi_1(R')$ may be realized by π_1–isomorphically embedding R and R' into the same once-punctured surface). We show that $\mathscr{P}$ and $\mathcal{CS}$ are quasi–isometric.

Define an equivariant map $\Phi \colon \mathscr{P} \to \mathcal{CS}$ by sending a rose R to the conjugacy class determined by one of its petals (the set of possible choices has diameter 1 by definition). Whenever vertices R and R' of $\mathscr{P}$ have a common petal we then have $d_{\mathcal{CS}}(\Phi(R), \Phi(R')) \le 2$, and whenever R and R' have a common cusp we have $d_{\mathcal{CS}}(\Phi(R), \Phi(R')) \le 1$ since in this case the petals of R and R' all correspond to simple closed curves on the same once-punctured surface S. Thus $\Phi \colon \mathscr{P} \to \mathcal{CS}$ is 2–Lipschitz. Similarly, let $\Psi \colon \mathcal{CS} \to \mathscr{P}$ be an equivariant map that sends each primitive conjugacy class α to any rose in which α appears as a petal. Since the set of such roses has diameter 1, Ψ is a coarse inverse for Φ. Hence, it suffices to show that Ψ is Lipschitz. For this, if α and β are adjacent vertices of $\mathcal{CS}$ we may choose a once-punctured surface S in which α and β are represented simple closed curves. Extending these curves to π_1–injectively embedded roses R and R' in S, we see that $d_{\mathscr{P}}(\Psi(\alpha), \Psi(\beta)) \le 3$ since R and R' have a common cusp and are respectively adjacent to $\Psi(\alpha)$ and $\Psi(\beta)$ by construction. This shows that Ψ is 3–Lipschitz and completes the proof. $\square$

Proof of Theorem 4.2. By Theorem 4.3, $\mathscr{I}$ is hyperbolic and the map $\mathcal{PL} \to \mathscr{I}$ is Lipschitz and coarsely alignment preserving. Since $\mathscr{I}$ and $\mathcal{CS}$ are $\mathrm{Out}(\mathbb{F})$–equivariantly quasi-isometric by Proposition 4.4, the same holds for the map $\mathbf{el}: \mathcal{PL} \to \mathcal{CS}$. The classification of the loxodromic elements of the action $\mathrm{Out}(\mathbb{F}) \curvearrowright \mathcal{CS}$ is proven in Corollary 4.13 below. $\qquad\square$

We also note that Bestvina and Feighn show the projection $\pi_{\mathcal{PL}}: \mathfrak{X} \to \mathcal{PL}$ is alignment preserving [BF2, Theorem 9.3]. By Lemma 3.1 it follows that the composition $\pi_{\mathcal{CS}} := \mathbf{el} \circ \pi_{\mathcal{PL}}: \mathfrak{X} \to \mathcal{CS}$ is alignment preserving. As for $\mathcal{F}$, for $G, H \in \mathfrak{X}$ we then define

$$d_{\mathcal{CS}}(G, H) = \mathrm{diam}_{\mathcal{CS}}(\pi_{\mathcal{CS}}(G) \cup \pi_{\mathcal{CS}}(H)).$$

Remark 4.5 (Historical context). In [KL1], Kapovich and Lustig use their intersection form (c.f. Theorem 2.7) to show that several free group analogs of the curve complex have infinite diameter. Among their proposed graphs were (up to quasi-isometry) versions of the free factor complex, free splitting complex, and what they call the intersection graph. Their definition of the intersection graph is the following: vertices are *all* conjugacy classes of very small trees and geodesic currents and a tree T is joined by and edge to a current μ if $\langle T, \mu \rangle = 0$. Although this graph is not connected (e.g. a free simplicial tree is isolated), the connected component containing the rational currents corresponding to primitive conjugacy classes is $\mathrm{Out}(\mathbb{F})$ invariant. This version of the intersection graph however is different from the graph $\mathscr{I}$ defined above, which is also referred to as the intersection graph in [Man]. The difference between these graphs lies in exactly *which* geodesic currents are allowed; different restrictions determine which electrification of $\mathcal{F}$ one obtains. Using $\mathcal{CS}$ avoids this ambiguity as well as having the added benefit of a natural and transparent definition.

4.1 Boundary of $\mathcal{CS}$

From Theorems 3.2 and 4.2 we deduce that $\partial \mathcal{CS} \cong \partial_{\mathcal{CS}} \mathcal{F}$. Our next lemmas show that $\partial_{\mathcal{CS}} \mathcal{F}$ is precisely the collection of classes of free arational trees in $\partial \mathcal{F}$. The first lemma follows easily from work of Coulbois–Hilion–Reynolds and Bestvina–Reynolds.

Lemma 4.6. *Let $T \in \partial \mathfrak{X}$ be free and arational and let μ be any geodesic current. If $\langle T, \mu \rangle = 0 = \langle S, \mu \rangle$ for some tree $S \in \mathfrak{X} \cup \partial \mathfrak{X}$, then S is also free and arational.*

Proof. By Theorem 2.7, the hypotheses imply $\mathrm{supp}(\mu) \subset L(T)$ and $\mathrm{supp}(\mu) \subset L(S)$. Using that T is free and arational, Theorem 2.8 moreover gives $\mathrm{supp}(\mu) = L'(T)$; thus $L'(T) \subset L(S)$. Now apply Proposition 4.2(i) and Corollary 4.3 of [BR] to conclude that $L(S) = L(T)$.

Recall from §2.3 that a tree is free if and only if its dual lamination does not contain $L(\alpha)$ for any nontrivial conjugacy class α. Since this necessarily holds for $L(T) = L(S)$, we conclude that S is free as well. Finally, if S were not arational then there would be a free factor A of $\mathbb{F}$ and an A–invariant subtree $S^A \subset S$ on which A acts with dense orbits. It would then follow that $L(T) = L(S)$ contains a leaf in $\partial^2 A \subset \partial^2 \mathbb{F}$. However, the fact that T is free and arational implies that $L(T)$ cannot contain a leaf in $\partial^2 A$ [Rey2, Lemma 2.1]. This contradiction shows that S must be arational and completes the proof of the lemma. $\qquad\square$

The following lemma is an application of a standard argument for showing that graphs which are similar to the curve graph of a surface have infinite diameter. See [Kob, MM, BR, KL1]. The details are provided for the reader's convenience.

Lemma 4.7. *Let $(G_i)_{i \geq 0}$ be a sequence of graphs in $\mathfrak{X}$ converging to a tree T in $\partial \mathfrak{X}$ which is free and arational. Then the projections $\pi_{\mathcal{CS}}(G_i)$ of G_i to $\mathcal{CS}$ are unbounded.*

Proof. Recall that $\partial \mathfrak{X}$ is realized as the boundary of $\mathfrak{X}$ inside the projective space $\mathbb{PR}^{\mathbb{F}}$. Thus the statement $G_i \to T$ means that there is a sequence of scaling constants $\lambda_i > 0$ such that $\langle \lambda_i G_i, c \rangle \to \langle T, c \rangle$ for every conjugacy class c in $\mathbb{F}$. Observe that that $d_{\mathfrak{X}}(G_0, G_i)$ is unbounded since $\pi_{\mathcal{F}}$ is Lipschitz and Theorem 2.9

implies that $\pi_{\mathcal{F}}(G_i)$ converges to the boundary of $\mathcal{F}$. Therefore, there is a conjugacy class c such that $\langle G_i, c \rangle$ is unbounded. If $\lambda_i \geq \varepsilon > 0$ for all i, this would imply

$$\langle T, c \rangle = \lim_i \langle \lambda_i G_i, c \rangle = \limsup_i \lambda_i \langle G_i, c \rangle \geq \varepsilon \limsup_i \langle G_i, c \rangle = \infty,$$

contradicting that $\langle T, c \rangle = \ell_T(c) < \infty$. Therefore, after passing to a subsequence, we may assume $\lambda_i \to 0$.

Let $\alpha_i \in \mathcal{CS}^0$ be a primitive loop in the projection $\pi_{\mathcal{CS}}(G_i)$; thus α_i corresponds to an embedded closed loop on G_i and so $\langle G_i, \alpha_i \rangle \leq 1$. Suppose that these curves do not go to infinity in $\mathcal{CS}$. Then, after passing to a subsequence and fixing some $x \in \mathcal{CS}^0$, we may assume that $d_{\mathcal{CS}}(x, \alpha_i) = M$ for all i. Build a geodesic $x = x_i^0, x_i^1, \ldots, x_i^M = \alpha_i$ for each $i \geq 0$. By definition of $\mathcal{CS}$, for each $0 \leq k \leq M-1$ there is a once-punctured surface S_i^k realizing the edge between x_i^k and x_i^{k+1}; let c_i^k be the conjugacy class corresponding to the cusp (i.e., peripheral curve) of S_i^k. Further, let R_i^k be the simplicial tree dual to the simple closed curve representing x_i^k on S_i^k, and for $1 \leq k \leq M$, let L_i^k be the simplicial tree dual to x_i^k on S_i^{k-1}. By construction

$$\langle R_i^k, x_i^k \rangle = \langle R_i^k, c_i^k \rangle = 0 \qquad \text{and} \qquad \langle L_i^k, x_i^k \rangle = \langle L_i^k, c_i^{k-1} \rangle = 0. \tag{4}$$

Now let $i \to \infty$ and, after passing to a subsequence, assume that everything converges projectively to either a tree or a geodesic current. Denote the limit by omitting the subscript. Since $\langle \cdot, \cdot \rangle$ is continuous and G_i converges projectively to the free arational tree T, we necessarily have $\langle T, \alpha \rangle = 0$, where α is a projective limit of $\alpha_i = x_i^M$ in $\mathbb{P}\mathrm{Curr}(\mathbb{F})$. To see this, choose scaling constants $\mu_i > 0$ so that $\mu_i \alpha_i$ converges to the current α. Letting $A > 0$ denote the length of the shortest loop in G_0, we have

$$\limsup_i \mu_i A \leq \limsup_i \mu_i \langle G_0, \alpha_i \rangle = \lim_i \langle G_0, \mu_i \alpha_i \rangle \to \langle G_0, \alpha \rangle < \infty.$$

Using $\langle G_i, \alpha_i \rangle \leq 1$, it follows that

$$\langle T, \alpha \rangle = \lim_i \langle \lambda_i G_i, \mu_i \alpha_i \rangle = \lim_i \lambda_i \mu_i \langle G_i, \alpha_i \rangle \leq \limsup_i \lambda_i \mu_i \leq (\limsup_i \mu_i)(\lim_i \lambda_i) = 0,$$

as claimed. By continuity of $\langle \cdot, \cdot \rangle$, we additionally have $\langle L^M, \alpha \rangle = 0$. Hence, L^M is free and arational by Lemma 4.6. Similarly we have $\langle L^M, c^{M-1} \rangle = 0 = \langle R^{M-1}, c^{M-1} \rangle$; thus R^{M-1} is also free and arational by Lemma 4.6. Using continuity and (4) again to pair R^{M-1} and L^{M-1} with x^{M-1}, we now see that L^{M-1} is free and arational as well. Applying this augment inductively, we conclude that R^0 is free and arational. This, however, contradicts the observation that $\langle R^0, x^0 \rangle = 0$ for the primitive conjugacy class $x^0 = x$ (recall that the sequence x_i^0 is constant). This shows that $d_{\mathcal{CS}}(x, \alpha_i) \to \infty$ as $i \to \infty$ and completes the proof. $\qquad\square$

Theorem 4.8 (Boundary of $\mathcal{CS}$). *The Gromov boundary $\partial \mathcal{CS}$ of the co-surface graph is* $\mathrm{Out}(\mathbb{F})$*–equivariantly homeomorphic to the subspace of $\partial \mathcal{F}$ consisting of classes of free arational trees.*

Proof. We use the alignment preserving map $\mathbf{el}' \colon \mathcal{F} \to \mathcal{CS}$ and Theorem 3.2 to identify $\partial \mathcal{CS} \cong \partial_{\mathcal{CS}} \mathcal{F}$. By Theorem 2.9 and Lemma 4.7, the set of free arational trees is contained in $\partial_{\mathcal{CS}} \mathcal{F}$. Further, if $T \in \partial \mathcal{F}$ then T is arational by Theorem 2.9. If T is not free, then by Theorem 2.2, T is dual to a measured lamination L on a once–punctured surface S. Let α_i be a sequence of nonseparating simple closed curves in S converging to the lamination L. Then α_i is also a sequence of rank 1 free factors in $\mathcal{F}$ converging to $T \in \partial \mathcal{F}$ with $\mathrm{diam}_{\mathcal{CS}} \mathbf{el}(\alpha_i) \leq 1$. Hence, $T \notin \partial_{\mathcal{CS}} \mathcal{F}$ and we conclude that the set of free arational trees equals $\partial_{\mathcal{CS}} \mathcal{F}$. $\qquad\square$

4.2 Quasi–isometric embeddings into $\mathcal{CS}$

We say that a finitely generated subgroup $\Gamma \leq \mathrm{Out}(\mathbb{F})$ **qi-embeds** into an $\mathrm{Out}(\mathbb{F})$ graph Y if some (equivalently any) orbit map $\Gamma \to Y$ is a quasi-isometric embedding. In this section, we prove that $\Gamma \leq \mathrm{Out}(\mathbb{F})$ qi-embeds into $\mathcal{CS}$ if and only if it is purely atoroidal and qi-embeds into $\mathcal{F}$. This answers Question 1.2 of I. Kapovich and clarifies the connection between the factor complex, the co-surface graph, and hyperbolic extensions of free groups.

Fix a rose $R \in \mathcal{X}$ and a primitive conjugacy class α represented by a petal of R. Fix a finitely generated subgroup $\Gamma \leq \mathrm{Out}(\mathbb{F})$ such that the orbit map $\Gamma \to \mathcal{F}$ given by $g \mapsto g \cdot \alpha$ is a quasi-isometric embedding. In [DT], we show that this implies that the orbit $\Gamma \cdot R$ has strong quasiconvexity properties in $\mathcal{X}$ (e.g. Theorem 2.6). For the application needed here, the following proposition from [DKT] is most convenient.

Proposition 4.9 (Folding rays to infinity [DKT, Proposition 5.6]). *Suppose that $\Gamma \leq \mathrm{Out}(\mathbb{F})$ is purely atoroidal and qi-embeds into $\mathcal{F}$. For any $k \geq 0$ there is a $K \geq 0$ such that if $(g_i)_{i \geq 0}$ is a k–quasigeodesic ray in Γ, then there is an infinite length folding ray $\gamma \colon \mathbf{I} \to \mathcal{X}$ parameterized at unit speed with the following properties:*

1. *The sets $\gamma(\mathbf{I})$ and $\{g_i R : i \geq 0\}$ have symmetric Hausdorff distance at most K.*

2. *The rescaled folding path $G_t = e^{-t} \cdot \gamma(t) \in cv$ converges to the arational tree $T \in \partial cv$ with the property that $\lim_{i \to \infty} g_i \cdot \alpha = \partial \pi_{\mathcal{F}}(T)$ in $\mathcal{F} \cup \partial \mathcal{F}$, where $\partial \pi_{\mathcal{F}}(T)$ is the projection of the projective class of T to the boundary of $\mathcal{F}$ (c.f. Theorem 2.9). Moreover, the action $\mathbb{F} \curvearrowright T$ is free.*

Given Proposition 4.9, we show that the orbit map from Γ to the co-surface graph is metrically proper.

Proposition 4.10. *Suppose that $\Gamma \leq \mathrm{Out}(\mathbb{F})$ is purely atoroidal and qi-embeds into the factor complex $\mathcal{F}$. Then for every $D \geq 0$ there is an $N \geq 0$ so that*

$$d_{\mathcal{CS}}(R, g \cdot R) \geq D$$

for all $g \in \Gamma$ with $|g| \geq N$.

Proof. Suppose not. Then there is a $D \geq 0$ and a sequence $h_i \in \Gamma$ with $d_{\mathcal{CS}}(R, h_i \cdot R) \leq D$ and $|h_i| \to \infty$ as $i \to \infty$. After passing to a subsequence, we may assume that $h_i \to q \in \partial \Gamma$. Since $\Gamma \to \mathcal{F}$ is a quasi-isometric embedding, there is a unique $\lambda \in \partial \mathcal{F}$ such that $h_i \cdot \alpha \to \lambda$ in $\mathcal{F} \cup \partial \mathcal{F}$. (Recall that $\alpha \in \pi_{\mathcal{F}}(R)$.)

Claim 4.11. *The sequence $(h_i)_{i \geq 0}$ can be replaced by a geodesic $(g_i)_{i \geq 0}$ in Γ such that $g_i \to q \in \partial \Gamma$ and $d_{\mathcal{CS}}(\alpha, g_i \cdot \alpha) \leq \overline{D}$, where $\overline{D} \geq 0$ depends only on the constant D and the orbit map $\Gamma \to \mathcal{F}$.*

Proof of claim. Let $(g_i)_{i \geq 0}$ be any geodesic sequence in Γ with $g_0 = 1$ and $g_i \to q$ in $\Gamma \cup \partial \Gamma$. Then for each $i \geq 0$ there is a $j \geq 0$ such that any geodesic $[0, h_j]$ passes within 2δ from g_i; thus the triple $(1, g_i, h_j)$ is 4δ–aligned in Γ. Since $\Gamma \to \mathcal{F}$ is a quasi-isometric embedding and $\mathcal{F}$ is hyperbolic, the stability of quasigeodesics (Proposition 2.1) implies that $\Gamma \to \mathcal{F}$ is alignment preserving. Therefore $\Gamma \to \mathcal{CS}$ is alignment preserving by Lemma 3.1 and so there is some $C \geq 0$ (depending only on the quasi-isometry constant of $\Gamma \to \mathcal{F}$) so that

$$d_{\mathcal{CS}}(\alpha, g_i \cdot \alpha) \leq d_{\mathcal{CS}}(\alpha, g_i \cdot \alpha) + d_{\mathcal{CS}}(g_i \cdot \alpha, h_j \cdot \alpha) \leq d_{\mathcal{CS}}(\alpha, h_j \cdot \alpha) + C \leq D + C =: \overline{D}. \qquad \square$$

Now let $(g_i)_{i \geq 0}$ be as in Claim 4.11 and apply Proposition 4.9 to obtain a folding ray $\gamma \colon \mathbf{I} \to \mathcal{X}$. By Proposition 4.9, the graphs $g_i \cdot R$ and $\gamma(t)$ both limit to the same free arational tree $T \in \partial \mathcal{X}$. Since $g_i \cdot R$ converges to T in $\overline{\mathcal{X}}$, Lemma 4.7 implies that $d_{\mathcal{CS}}(R, g_i \cdot R) \to \infty$, contradicting the construction of $(g_i)_{i \geq 0}$. $\square$

Theorem 4.12 (Qi-embedding into $\mathcal{CS}$). *Let Γ be a finitely generated subgroup of $\mathrm{Out}(\mathbb{F})$. Then Γ qi-embeds into the co-surface graph $\mathcal{CS}$ if and only if Γ is purely atoroidal and convex cocompact.*

Proof. First, if Γ qi-embeds into $\mathcal{CS}$, then any orbit map $\Gamma \to \mathcal{F}$ is a quasi-isometric embedding since the $\mathrm{Out}(\mathbb{F})$–equivariant map $\mathbf{el}' \colon \mathcal{F} \to \mathcal{CS}$ is Lipschitz. Let $g \in \Gamma$ be infinite order. Since no power of g fixes a vertex of $\mathcal{F}$, g is fully irreducible. If g were not atoroidal, then Bestvina–Handel [BH1] proved that g is induced by a pseudo–Anosov homeomorphism on a once-punctured surface. However, if this were the case, then g would have bounded orbits in $\mathcal{CS}$, a contradiction.

Now suppose that Γ is purely atoroidal and that the orbit map $\Gamma \to \mathcal{F}$ given by $g \mapsto g \cdot \alpha$, for some rank 1 free factor α, is a quasi-isometric embedding. Since $\mathbf{el}' \colon \mathcal{F} \to \mathcal{CS}$ is alignment preserving, it follows that the map $g \mapsto g \cdot \mathbf{el}'(\alpha) \in \mathcal{CS}$ is alignment preserving. Proposition 4.10 shows that this orbit map $\Gamma \to \mathcal{CS}$ is also metrically proper. We thus conclude that $\Gamma \to \mathcal{CS}$ is a quasi-isometric embedding by Lemma 3.4 $\square$

Corollary 4.13. *An element of* $\mathrm{Out}(\mathbb{F})$ *acts loxodromically on* $\mathcal{CS}$ *if and only if it is fully irreducible and atoroidal.*

Proof. An element g of $\mathrm{Out}(\mathbb{F})$ acts loxodromically on $\mathcal{CS}$ if and only if $\langle g \rangle$ qi-embeds into $\mathcal{CS}$. By Theorem 4.12, this is the case if and only if g is atoroidal and acts loxodromically on $\mathcal{F}$. Since the loxodromic elements for the action $\mathrm{Out}(\mathbb{F}) \curvearrowright \mathcal{F}$ are precisely the fully irreducible elements [BF2, Theorem 9.3], the corollary follows. $\square$

5 Lifting to covers

In this section we show that orbits in the co-surface graph are well-behaved when passing to finite index subgroups of $\mathbb{F}$. Specifically, we show that if $\Gamma \leq \mathrm{Out}(\mathbb{F})$ qi-embeds into $\mathcal{CS}$, then for any finite index $H \leq \mathbb{F}$, the induced subgroup Γ^H of $\mathrm{Out}(H)$ qi-embeds into the co-surface graph of H. This proposition will be necessary in §7, but it may also be of independent interest.

Fix a finite index subgroup H of $\mathbb{F}$ and let $\Gamma \leq \mathrm{Out}(\mathbb{F})$ be finitely generated. Denote by Γ_H the subgroup of Γ consisting of outer automorphisms which fix the conjugacy class of H, and let Γ^H be the induced subgroup of $\mathrm{Out}(H)$. That is, $f \in \Gamma^H$ if there is an automorphism ϕ of $\mathbb{F}$ whose outer class is in Γ such that $\phi|_H$ is in the outer class f. These groups fit into a short exact sequence:

Lemma 5.1. *For any* $\Gamma \leq \mathrm{Out}(\mathbb{F})$ *and* $H \leq \mathbb{F}$ *finite index, there is a short exact sequence*

$$1 \longrightarrow N(H)/H \longrightarrow \Gamma^H \longrightarrow \Gamma_H \longrightarrow 1,$$

where $N(H) \leq \mathbb{F}$ *is the normalizer of* H *and* $N(H)/H$ *is finite.*

Proof. Let $\mathrm{Aut}_H(\mathbb{F})$ be the subgroup of $\mathrm{Aut}(\mathbb{F})$ fixing H, and let $\mathrm{Aut}^{\mathbb{F}}(H)$ be the image of $\mathrm{Aut}_H(\mathbb{F})$ under the restriction map $\mathrm{res}\colon \mathrm{Aut}_H(\mathbb{F}) \to \mathrm{Aut}(H)$. Let $\mathrm{Out}_H(\mathbb{F})$ and $\mathrm{Out}^{\mathbb{F}}(H)$ denote the images of these subgroups under the projection homomorphisms $p_{\mathbb{F}}\colon \mathrm{Aut}(\mathbb{F}) \to \mathrm{Out}(\mathbb{F})$ and $p_H\colon \mathrm{Aut}(H) \to \mathrm{Out}(H)$, respectively. These fit into the following diagram:

$$\mathrm{Aut}^{\mathbb{F}}(H) \xleftarrow{\quad\mathrm{res}\quad} \mathrm{Aut}_H(\mathbb{F})$$
$$\hat{\pi} = p_{\mathbb{F}} \circ \mathrm{res}^{-1}$$
$$\Big\downarrow{p_H} \qquad\qquad \Big\downarrow{p_{\mathbb{F}}}$$
$$1 \longrightarrow N(H)/H \longrightarrow \mathrm{Out}^{\mathbb{F}}(H) \xrightarrow{\quad\pi\quad} \mathrm{Out}_H(\mathbb{F}) \longrightarrow 1$$

The restriction map $\mathrm{res}\colon \mathrm{Aut}_H(\mathbb{F}) \to \mathrm{Aut}^{\mathbb{F}}(H)$ is surjective by definition of $\mathrm{Aut}^{\mathbb{F}}(H)$. It is also injectve. To see this, suppose $\phi_1, \phi_2 \in \mathrm{Aut}_H(\mathbb{F})$ are automorphisms fixing H that have the same restriction to H, and let $a \in \mathbb{F}$ be arbitrary. We may choose $k > 0$ so that $a^k \in H$ and consequently $\phi_1(a)^k = \phi_2(a)^k$. As elements of $\mathbb{F}$ have unique roots, it follows that $\phi_1(a) = \phi_2(a)$, proving the claim.

Identify $\mathbb{F}$ with is image in $\mathrm{Aut}(\mathbb{F})$ under the map sending each $a \in \mathbb{F}$ to the conjugation automorphism $g \mapsto aga^{-1}$. With respect to this identificaion, the normalizer of H in $\mathbb{F}$ is exactly

$$N(H) = \mathbb{F} \cap \mathrm{Aut}_H(\mathbb{F}) = \ker(p_{\mathbb{F}}\colon \mathrm{Aut}_H(\mathbb{F}) \to \mathrm{Out}_H(\mathbb{F})).$$

Thus the kernel of the composition $\hat{\pi} := p_{\mathbb{F}} \circ \mathrm{res}^{-1}\colon \mathrm{Aut}^{\mathbb{F}}(H) \to \mathrm{Out}_H(\mathbb{F})$ may be naturally identified with $N(H)$. Observe that every inner automorphism of H lifts to an inner automorphism of $\mathbb{F}$ via the map res^{-1}. Since the resulting inner automorphism of $\mathbb{F}$ clearly normalizes H, we furthermore see that $H \cong \mathrm{Inn}(H) = \ker(p_H)$ is contained in $N(H) = \ker(\hat{\pi})$. It follows that the homomorphism $\hat{\pi}$ descends to a well-defined epimorphism $\pi\colon \mathrm{Out}^{\mathbb{F}}(H) \to \mathrm{Out}_H(\mathbb{F})$ with kernel $\ker(\pi) = \ker(\hat{\pi})/\ker(p_H) \cong N(H)/N$. The lemma now follows by observing that for any subgroup $\Gamma \leq \mathrm{Out}(\mathbb{F})$, we have $\Gamma_H = \Gamma \cap \mathrm{Out}_H(\mathbb{F})$ and $\Gamma^H = \pi^{-1}(\Gamma_H)$. $\square$

The main result of this section is the following:

Proposition 5.2. *Let H be a finite index subgroup of $\mathbb{F}$ and let Γ^H denote the subgroup of $\mathrm{Out}(H)$ induced by elements of $\Gamma \leq \mathrm{Out}(\mathbb{F})$ that stabilize the conjugacy class of H. If Γ is finitely generated and qi-embeds into $\mathcal{CS}$, then Γ^H also qi-embeds into $\mathcal{CS}(H)$.*

Let us briefly remark on the use of the co-surface graph in the statement of Proposition 5.2. In particular, the corresponding statement for the factor graph $\mathcal{F}$ is false. For example, let $\phi \in \mathrm{Out}(\mathbb{F})$ be an automorphism that can be represented by a pseudo-Anosov on a once–punctured surface S. Let H be a subgroup of $\mathbb{F} \cong \pi_1(S)$ corresponding to a cover $\tilde{S} \to S$ with at least 2 punctures. The cyclic subgroup $\Gamma = \langle \phi \rangle$ then quasi-isometrically embeds into $\mathcal{F}(\mathbb{F})$ since ϕ is fully irreducible. However, Γ^H does not qi-embed into $\mathcal{F}(H)$. Indeed, Γ^H is a virtually cyclic group whose infinite order elements are represented by lifts of powers of ϕ; since each such lift permutes the punctures of $\tilde{S}$ and each puncture represents a rank 1 free factor of H, Γ^H has bounded orbits in $\mathcal{F}(H)$. This suggests that $\mathcal{CS}$ is a better tool for studying finite index subgroups of $\mathbb{F}$.

The proof of Proposition 5.2 requires the following result of Reynolds whose proof uses ideas of Guirardel.

Lemma 5.3 (Reynolds [Rey1, Lemma 4.2])**.** *Suppose that $G \curvearrowright T$ is an indecomposable action and that $H \leq G$ is finitely generated and finite index. Then the action $H \curvearrowright T^H$ is indecomposable.*

We will use the lemma in the following form.

Corollary 5.4. *Suppose that $T \in \partial \mathcal{X}$ is free and arational and that $H \leq \mathbb{F}$ is finite index. Then the minimal H–subtree T^H is also free and arational.*

Proof. Clearly, T^H is free because T is free. Since T is arational, it is indecomposable (Theorem 2.2); thus T^H is also indecomposable by Lemma 5.3. Using Theorem 2.2 again, we conclude that T^H is arational. $\square$

5.1 The Outer space of a subgroup

Fix $H \leq \mathbb{F}$ a subgroup of finite index $[\mathbb{F} : H] = n$. Then H is a free group of rank $1 - n(1 - \mathrm{rk}(\mathbb{F}))$ and we denote its Outer space by $\mathcal{X}(H)$. Recalling that $\mathcal{R}$ is our fixed $\mathrm{rk}(\mathbb{F})$–petal rose used to mark graphs of $\mathcal{X}$, we let $H|\mathcal{R}$ denote the corresponding H–cover and fix a homotopy equivalence between $H|\mathcal{R}$ and a rose $\mathcal{R}_H$.

There is a natural inclusion $i^* \colon \mathcal{X} \to \mathcal{X}(H)$ defined by taking H–covers and lifting markings. In details, if $\phi \colon \mathcal{R} \to G$ is a marked metric $\mathbb{F}$–graph, then the H–cover $H|G$ is a metric H–graph of volume n and we may choose a lift $\phi_H \colon H|\mathcal{R} \to H|G$. Any other lift $\tilde{\phi} \colon H|\mathcal{R} \to H|G$ is then obtained by precomposing ϕ_H with an element of the deck group of $H|\mathcal{R} \to \mathcal{R}$, which is isomorphic to $N(H)/H$. Since for each such deck transformation $d \in N(H)/H$, there is a graph isometry ρ_d of $H|G$ such that $\rho_d \circ \phi_H \sim \phi_H \circ d$, we see that the equivalence class of $(H|G, \phi_H)$ in $\mathcal{X}(H)$ is well-defined. Using the homotopy lifting property, we additionally see that this induces a well-defined map

$$i^* \colon \mathcal{X} \to \mathcal{X}(H)$$

given by $[G, \phi] \mapsto [H|G, \phi_H]$. Also note that if $\tilde{g} \in \Gamma^H$ projects to $g \in \Gamma_H$, then $\tilde{g} \cdot i^*(G) = i^*(g \cdot G)$. Even better, this map is an isometry with respect to the Lipschitz metric:

Proposition 5.5. *Let $H \leq \mathbb{F}$ be a finite index subgroup. Then the induced map $i^* \colon \mathcal{X} \to \mathcal{X}(H)$ is an isometry with respect to the Lipschitz metric. Moreover, i^* maps folding paths in $\mathcal{X}$ to folding paths in $\mathcal{X}(H)$.*

Proof. Fix G_1 and G_2 in $\mathcal{X}$. If $f \colon G_1 \to G_2$ is an optimal change of marking, then choosing a lift $f_H \colon H|G_1 \to H|G_2$ we see that since $\mathrm{Lip}(f_H) = \mathrm{Lip}(f)$,

$$d_{\mathcal{X}(H)}(i^*(G_1), i^*(G_2)) = d_{\mathcal{X}(H)}(H|G_1, H|G_2) \leq d_{\mathcal{X}}(G_1, G_2).$$

Further,

$$e^{d_{\mathcal{X}(H)}(H|G_1, H|G_2)} = \sup_{h \in H} \frac{\ell(h|(H|G_2))}{\ell(h|(H|G_1))} = \sup_{h \in H} \frac{\ell(h|G_2)}{\ell(h|G_1)}.$$

Let $a \in \mathbb{F}$ be such that a is optimally stretched by $f \colon G_1 \to G_2$, i.e. $\ell(a|G_2) = \mathrm{Lip}(f)\ell(a|G_1)$, and let $k > 0$ be the smallest positive integer such that $a^k \in H$. Then using our observation above

$$e^{d_{\mathcal{X}}(G_1, G_2)} = \frac{\ell(a|G_2)}{\ell(a|G_1)} = \frac{\ell(a^k|G_2)}{\ell(a^k|G_1)} \leq \sup_{h \in H} \frac{\ell(h|G_2)}{\ell(h|G_1)} = e^{d_{\mathcal{X}(H)}(H|G_1, H|G_2)}.$$

Hence, we also have $d_{\mathcal{X}}(G_1, G_2) \leq d_{\mathcal{X}(H)}(i^*(G_1), i^*(G_2))$ showing that $i^* \colon \mathcal{X} \to \mathcal{X}(H)$ is an isometry.

To show that i^* maps folding paths to folding paths, let $\{G_t\}_{t \in [0,L]}$ be a folding path in $\mathcal{X}$ with corresponding folding maps $\{\phi_{st} \colon G_s \to G_t\}_{s<t}$ (see §2.5 for the definitions of folding maps and folding paths). Set $\tilde{G}_t = H|G_t = i^*(G_t)$ with covering map $p_t \colon \tilde{G}_t \to G_t$. We claim that $\{\tilde{G}_t\}$ is a folding path in $\mathcal{X}(H)$ with corresponding lifted folding maps $\{\tilde{\phi}_{st} \colon H|G_s \to H|G_t\}_{s<t}$. Indeed, that each of these lifts is a local $\mathrm{Lip}(\tilde{\phi}_{st})$–homothety with $\mathrm{Lip}(\tilde{\phi}_{st}) = d_{\mathcal{X}(H)}(\tilde{G}_s, \tilde{G}_t)$ follows from what we have already shown. What's more, the equality $p_t \circ \tilde{\phi}_{st} = \phi_{st} \circ p_s$ implies that the illegal turn structure that $\tilde{\phi}_{st}$ induces on $\tilde{G}_s$ is exactly the lift (via p_s) of the illegal turn structure that ϕ_{st} induces on G_s. From this it follows that each $\tilde{\phi}_{st}$ is a folding map and, by the uniqueness of folding paths [BF2, §2], that $\{\tilde{G}_t\}$ is the folding path determined by $\tilde{\phi}_{0L} \colon \tilde{G}_0 \to \tilde{G}_L$. $\quad\square$

5.2 Proof of Proposition 5.2

Combining our work in the previous sections, we now turn to the proof of Proposition 5.2. Let $R \in \mathcal{X}$ be a rose with a petal representing $\alpha \in \mathcal{CS}^0$. Let $\tilde{\alpha} \in \mathcal{CS}^0(H)$ be primitive conjugacy class of H represented by an embedded loop of $H|R$ covering this petal of R. Note that $\alpha \in \pi_{\mathcal{CS}}(R)$ and $\tilde{\alpha} \in \pi_{\mathcal{CS}}(H|R)$, where $\pi_{\mathcal{CS}}$ is the projection from outer space to the corresponding co-surface graph.

Fix a finite generating set $\tilde{S}$ of Γ^H with projection $S \subset \Gamma_H$. By abuse of notation, we identify Γ^H and Γ_H with the corresponding Cayley graphs $\mathrm{Cay}(\Gamma^H, S)$ and $\mathrm{Cay}(\Gamma_H, \tilde{S})$, which are geodesic metric spaces. The orbits maps $\tilde{g} \mapsto \tilde{g} \cdot \tilde{\alpha}$ and $\tilde{g} \mapsto \tilde{g} \cdot H|R$ then extend equivariantly to continuous maps $\psi \colon \Gamma^H \to \mathcal{CS}(H)$ and $\Psi \colon \Gamma^H \to \mathcal{X}(H)$, and similarly for Γ_H. Note also that the projection $\Gamma^H \to \Gamma_H$ is a quasi-isometry.

Proof of Proposition 5.2. By Lemma 3.4, it suffices to show that the orbit map $\psi \colon \Gamma^H \to \mathcal{CS}(H)$ is alignment preserving and metrically proper. The strategy is to relate the corresponding orbit in $\mathcal{X}$ with folding paths via Theorem 2.6 and to use that these folding paths lift to $\mathcal{X}(H)$ by Proposition 5.5. Note that since Γ qi-embeds into $\mathcal{CS}$, Γ is hyperbolic. Further, as Γ_H is finite index in Γ, Γ_H, and hence Γ^H, is also hyperbolic.

In details, let $(\tilde{g}_-, \tilde{g}_0, \tilde{g}_+)$ be an aligned triple in Γ^H and let $\tilde{\gamma} \colon \mathbf{I} \to \Gamma^H$ be a geodesic passing through $\tilde{g}_0$ with $\tilde{\gamma}(\mathbf{I}_\pm) = \tilde{g}_\pm$. The projection $\gamma \colon \mathbf{I} \to \Gamma_H \leq \Gamma$ is then a uniform quasigeodesic and so maps to a uniform quasigeodesic in $\mathcal{CS}$ by assumption. Hence Theorem 2.6 provides a folding path $\{G_t\}_{t \in [0,L]}$ in $\mathcal{X}$ that has uniformly bounded symmetric Hausdorff distance from the image of the composition $\mathbf{I} \to \Gamma_H \to \mathcal{X}$. Since $i^* \colon \mathcal{X} \to \mathcal{X}(H)$ is a Γ^H–equivariant isometric embedding (Proposition 5.5), the same holds for the image of $\mathbf{I} \to \Gamma^H \to \mathcal{X}(H)$ and the folding path $i^*(G_t) = H|G_t$. In particular, the three points $\tilde{g}_- \cdot H|R$, $\tilde{g}_0 \cdot H|R$, and $\tilde{g}_+ \cdot H|R$ all lie within uniformly bounded symmetric Hausdorff distance of the geodesic $\{H|G_t\}$ in $\mathcal{X}(H)$. Since the projection $\pi_{\mathcal{CS}} \colon \mathcal{X}(H) \to \mathcal{CS}(H)$ is alignment preserving, it follows that the triple $(\tilde{g}_- \cdot \tilde{\alpha}, \tilde{g}_0 \cdot \tilde{\alpha}, \tilde{g}_+ \cdot \tilde{\alpha})$ is uniformly aligned in $\mathcal{CS}(H)$. Therefore the orbit map $\Gamma^H \to \mathcal{CS}(H)$ is alignment preserving.

It remains to show that $\Gamma^H \to \mathcal{CS}(H)$ is metrically proper. If this were not the case, then there is a sequence of $x_i \in \Gamma^H$ with $|x_i| \to \infty$ but $d_{\mathcal{CS}(H)}(\tilde{\alpha}, x_i \cdot \tilde{\alpha}) \leq D$ for some $D \geq 0$. By compactness of $\Gamma^H \cup \partial \Gamma^H$, after passing to a subsequence we may assume that $x_i \to q \in \partial \Gamma^H = \partial \Gamma_H = \partial \Gamma$. Since we have already shown that $\Gamma^H \to \mathcal{CS}(H)$ is coarsely alignment preserving, use Claim 4.11 to obtain a geodesic $(\tilde{g}_i)_{i \geq 0}$ in Γ^H such that $\tilde{g}_i \to q$ in $\Gamma^H \cup \partial \Gamma^H$ but $d_{\mathcal{CS}(H)}(\tilde{\alpha}, \tilde{g}_i \cdot \tilde{\alpha})$ is uniformly bounded.

Now pass to a subsequence such that $g_i \cdot R$ converges to T in $\overline{\mathcal{X}}$. Since $(g_i \cdot \alpha)_{i \geq 0}$ is a quasigeodesic in $\mathcal{CS}$, we see combining Theorem 2.9 and Theorem 4.8 that the tree T is free and arational and that $g_i \cdot \alpha$ converges to the equivalence class of T in $\mathcal{CS} \cup \partial \mathcal{CS}$. Hence, $(H|g_i \cdot R = \tilde{g}_i \cdot H|R)_{i \geq 0}$ converges in $\overline{\mathcal{X}}(H)$ to the tree T^H. By Corollary 5.4, T^H is free and arational. Hence, by Lemma 4.7, $\pi_{\mathcal{CS}(H)}(\tilde{g}_i \cdot H|R) = \tilde{g}_i \cdot \tilde{\alpha}$ is unbounded in $\mathcal{CS}(H)$. In particular, $d_{\mathcal{CS}(H)}(\tilde{\alpha}, \tilde{g}_i \cdot \tilde{\alpha}) \to \infty$ contradicting the conclusion of the previous paragraph. This shows that $\Gamma^H \to \mathcal{CS}(H)$ is metrically proper and completes the proof. $\quad\square$

6 Flaring of simple conjugacy classes in Outer space orbits

In this section, we use the geometry of Outer space and the nature of folding paths to analyze how the lengths of conjugacy classes behave along the orbit $\Gamma \cdot R \subset \mathcal{X}$ of certain subgroups $\Gamma \leq \mathrm{Out}(\mathbb{F})$. When Γ qi-embeds into the factor complex $\mathcal{F}$, we will find that for simple conjugacy classes $\alpha \in \mathbb{F}$, the length $\ell(\alpha | g \cdot R)$ grows roughly exponentially in the distance from a certain uniformly bounded-diameter subset $\rho_\Gamma^R(\alpha) \subset \Gamma$. Our analysis culminates in the rather technical Lemma 6.10 which establishes this exponential flaring not only for simple classes α, but also for all conjugacy classes β that are, in a sense, "well-aligned" with α at the points of $\rho_\Gamma^R(\alpha) \cdot R$ (see Definition 6.6). Lemma 6.10 moreover shows that this exponential growth is *uniform* in all such α and β.

6.1 Uniform bounded backtracking

It is well known that any map $f \colon G \to H$ of metric core graphs has **bounded backtracking**, meaning that there is a constant $\mathrm{BBT}(f) \geq 0$ such that for any two points $p, q \in \tilde{G}$ in the universal cover and any lift $\tilde{f} \colon \tilde{G} \to \tilde{H}$ of f one has that the path $\tilde{f}([p, q])$ is contained in the $\mathrm{BBT}(f)$–neighborhood of the geodesic segment $[\tilde{f}(p), \tilde{f}(q)]$; see, e.g., [GJLL] or [CHL]. We will need a uniform bound on the constant $\mathrm{BBT}(f)$ over a broad family of graph maps. While bounds of this type are certainly well known to experts (see, e.g., [BFH, Lemma 3.1]), we include a short proof here for completeness.

Lemma 6.1 (Backtracking bound). *For every $D > 0$ there exists a constant $C > 0$ so that if $G, H \in \mathcal{X}$ satisfy $d_{\mathcal{X}}^{\mathrm{sym}}(G, H) \leq D$, then there exists a change of marking map $\phi \colon G \to H$ with $\mathrm{BBT}(\phi) \leq C$.*

Proof. The hypothesis ensures there are maps $\phi \colon G \to H$ and $\varphi \colon H \to G$ so that $\phi \circ \varphi \simeq \mathrm{Id}_H$, $\varphi \circ \phi \simeq \mathrm{Id}_G$, and $\mathrm{Lip}(\phi), \mathrm{Lip}(\varphi) \leq e^D$. Set $K = e^D$. Choose any lifts $\tilde{\phi} \colon \tilde{G} \to \tilde{H}$ and $\tilde{\varphi}_0 \colon \tilde{H} \to \tilde{G}$ of these maps to the universal covers, and note that $\tilde{\phi}$ and $\tilde{\varphi}_0$ are both equivariant with respect to the $\mathbb{F}$ actions on the trees $\tilde{G}$ and $\tilde{H}$. Note also that $\mathrm{Lip}(\tilde{\phi}), \mathrm{Lip}(\tilde{\varphi}_0) \leq K$.

Choose a basepoint $v \in \tilde{G}$ and let $w = \tilde{\phi}(v) \in \tilde{H}$. Since the orbit $\mathbb{F} \cdot v$ is 1–dense in $\tilde{G}$, we may choose $a \in \mathbb{F}$ so that $d_{\tilde{G}}(v, a \cdot \tilde{\varphi}_0(w)) \leq 1$. Define a new map $\tilde{\varphi} \colon \tilde{H} \to \tilde{G}$ by the rule $\tilde{\varphi}(q) = a \cdot \tilde{\varphi}_0(q)$ and note that we still have $\mathrm{Lip}(\tilde{\varphi}) \leq K$. We now have that $d_{\tilde{G}}(v, \tilde{\varphi} \circ \tilde{\phi}(v)) \leq 1$. By equivariance and the fact that $\tilde{\varphi}$ and $\tilde{\phi}$ are K–Lipschitz, it easily follows that $\tilde{\varphi} \circ \tilde{\phi}$ moves points at most distance $K^2 + 2$. From this, and the inequalities $\mathrm{Lip}(\tilde{\varphi}), \mathrm{Lip}(\tilde{\phi}) \leq K$, one may conclude that $\tilde{\phi}$ is a $(2K^2 + 4)$–quasi-isometry. Thus for any geodesic segment $[p, q] \subset \tilde{G}$, the image $\tilde{\phi}([p, q])$ is a $(2K^2 + 4)$–quasigeodesic in $\tilde{H}$. It now follows from Proposition 2.1 that $\tilde{\phi}([p, q])$ is contained in the $R_0(2e^{2D} + 4, 0)$–neighborhood of $[\tilde{\phi}(p), \tilde{\phi}(q)]$, as required. $\square$

If $\varphi \colon T_0 \to T$ is an $\mathbb{F}$–equivariant map between free simplicial trees, we also write $\mathrm{BBT}(\varphi)$ for the bounded backtracking constant of the induced map $T_0 / \mathbb{F} \to T / \mathbb{F}$ of the quotient graphs.

6.2 Bestvina–Feighn folding

Here, we recall some additional facts about folding paths that we will need in the proof of Lemma 6.10. Recalling the notation from §2.5, we see that the folding maps $\{\phi_{st} \colon G_s \to G_t\}_{s < t}$ associated to a folding path $\gamma \colon \mathbf{I} \to \mathcal{X}$ give rise to a well-defined illegal turn structure on each graph $G_s = \gamma(s)$ in the image (except for the right endpoint $\gamma(\mathbf{I}_+)$ when $\mathbf{I}_+ < \infty$). We then say that an immersed path (i.e., segment) in G_s is **legal** if it only takes legal turns; notice that for any legal path $\beta \colon \mathbf{J} \to G_s$, the composition $\phi_{st} \circ \beta$ is a legal path in G_t. An immersed path in G_s will be called **illegal** if it does not contain a legal subpath of length 3.

If $G_t \in \mathcal{X}$ is equipped with an illegal turn structure, then for every conjugacy class α of $\mathbb{F}$ the immersed loop $\alpha | G_t \to G_t$ breaks into maximal legal segments separated by illegal turns. Following our convention from [DT, §6], the **legal length** of $\alpha | G_t$ is defined to be the sum $\mathrm{leg}(\alpha | G_t)$ of the lengths of those maximal legal segments that have length at least 3. The following basic fact appears as Lemma 6.10 of [DT] and follows directly from work in [BF2]:

Lemma 6.2 ([DT]). *For any folding path G_t, $t \in [a, b]$, every nontrivial conjugacy class $\alpha \in \mathbb{F}$ satisfies*

$$\mathrm{leg}(\alpha | G_b) \geq \mathrm{leg}(\alpha | G_a) \left(\frac{1}{3} \right) e^{b-a}.$$

We also have the following technical result of Bestvina–Feighn; for the statement, recall that a nontrivial element $a \in \mathbb{F}$ (or the conjugacy class α thereof) is **simple** if a is contained in a proper free factor of $\mathbb{F}$.

Lemma 6.3 (Bestvina–Feighn [BF2, Lemma 5.8]). *There exists a constant B_1 depending only on $\mathrm{rk}(\mathbb{F})$ with the following property. If G_t, $t \in [a, b]$ is a folding path and α is simple with $\alpha | G_t$ illegal for all $t \in [a, b]$, then either $\ell(\alpha | G_a) > 2\ell(\alpha | G_b)$ or else $d_{\mathcal{F}}(G_a, G_b) \leq B_1$.*

For any subgroup $A \leq \mathbb{F}$ (or conjugacy class α) the illegal turn structure on G_t pulls back to give an illegal turn structure on $A | G_t$ (or $\alpha | G_t$). This gives a notion of legal and illegal paths in $A | G_t$: an immersed path in $A | G_t$ is (il)legal if and only if it is mapped to an (il)legal path in G_t. Define now the **illegality constant**

$$\mathfrak{I} := (2\,\mathrm{rk}(\mathbb{F}) - 1)(18\breve{m}(3\,\mathrm{rk}(\mathbb{F}) - 3) + 6),$$

where $\breve{m}$ is the maximal possible number of illegal turns in any illegal turn structure on any graph $G \in \mathcal{X}$ (so $\breve{m}$ is linear in $\mathrm{rk}(\mathbb{F})$). In [BF2], Bestvina and Feighn introduced the following projections to folding paths.

Definition 6.4. Given a folding path $\gamma \colon \mathbf{I} \to \mathcal{X}$ and a free factor A of $\mathbb{F}$, set

$$\mathrm{left}_\gamma(A) = \inf\{t \in \mathbf{I} : A | \gamma(t) \text{ has an immersed legal segment of length } 3\} \quad \text{and}$$
$$\mathrm{right}_\gamma(A) = \sup\{t \in \mathbf{I} : A | \gamma(t) \text{ has an immersed illegal segment of length } \mathfrak{I}\}.$$

We similarly define $\mathrm{left}_\gamma(\alpha)$ and $\mathrm{right}_\gamma(\alpha)$ for every nontrivial conjugacy class α of $\mathbb{F}$. Note that these definitions agree when α is a primitive conjugacy class.

The following technical result played a key roll in Bestvina and Feighn's proof of Theorem 2.3.

Proposition 6.5 (Bestvina–Feighn [BF2, Proposition 6.10]). *There exists a universal constant B_2 depending only on $\mathrm{rk}(\mathbb{F})$ such that for every folding path $\gamma \colon \mathbf{I} \to \mathcal{X}$ and free factor A of $\mathbb{F}$ we have*

$$d_{\mathcal{F}}\big(\gamma(\mathrm{left}_\gamma(A)), \gamma(\mathrm{right}_\gamma(A))\big) \leq B_2.$$

We henceforth write B for the universal constant $\max\{B_1, B_2\}$, where B_1, B_2 are the constants provided by Lemma 6.3 and Proposition 6.5.

6.3 Flaring and almost containment

The following notion will help us relate distances in the metric graph bundle $\mathscr{E}$ (see §7.1) to conjugacy lengths along folding paths.

Definition 6.6. Suppose that α, β are nontrivial conjugacy classes of $\mathbb{F}$ and that $G \in \mathcal{X}$ is a point in Outer space with corresponding $\mathbb{R}$–tree $\tilde{G}$. We say that β **is k–almost contained in α at G**, for $k \geq 0$, if there exists an axis $\mathscr{A} \subset \tilde{G}$ for (an element of) α and a fundamental domain B of an axis for (an element of) β in $\tilde{G}$ so that $B \setminus \mathscr{A}$ is a (possibly degenerate) connected segment of length at most k. This is to say that B is contained in $\mathscr{A}$ except for a subsegment of length at most k.

Equivalently, β is k–almost contained in α at G if and only if there exist axes $\mathscr{A}, \mathscr{B} \subset \tilde{G}$ for elements of the classes α and β, respectively, so that $\ell(\mathscr{A} \cap \mathscr{B}) \geq \ell(\beta | G) - k$. We also observe that α is always 0–contained in itself.

Lemma 6.7. *For any $k \geq 0$ and $D \geq 0$ there exists a constant $k' \geq 0$ so that if β is k–almost contained in α at $G \in \mathcal{X}$, then β is k'–almost contained in α at H for any $H \in \mathcal{X}$ with $d_{\mathcal{X}}^{\mathrm{sym}}(G, H) \leq D$.*

Proof. By assumption, we may may choose elements $a \in \alpha$ and $b \in \beta$ whose respective axes $\mathscr{A}, \mathscr{B} \subset \tilde{G}$ satisfy $\ell(\mathscr{A} \cap \mathscr{B}) \geq \ell(\beta|G) - k$. Orient $\mathscr{B}$ so that b translates in the forward direction, and let p and q be the initial and terminal endpoints of the segment $\mathscr{A} \cap \mathscr{B} \subset \mathscr{B}$. Also let $x = b \cdot p \in \mathscr{B}$. By k–almost containment, we have that either $d_{\tilde{G}}(x, q) \leq k$ or else $x \in \mathscr{A} \cap \mathscr{B}$.

By [Lemma 6.1](#), there is an e^D–Lipschitz change of marking map $\varphi \colon G \to H$ with $\mathrm{BBT}(\varphi) \leq C$ for some $C > 0$ depending only on D. Choose a lift $\tilde{\varphi} \colon \tilde{G} \to \tilde{H}$ and let μ be the (possibly empty) geodesic segment obtained by removing the length–C initial and terminal segments of $[\tilde{\varphi}(p), \tilde{\varphi}(q)]$. Let $\mathscr{A}', \mathscr{B}' \subset \tilde{H}$ be the axes for $a \curvearrowright \tilde{H}$ and $b \curvearrowright \tilde{H}$. By definition of bounded backtracking, $\tilde{\varphi}(\mathscr{A})$ is contained in the C–neighborhood of $\mathscr{A}'$; in particular $\tilde{\varphi}(p)$ and $\tilde{\varphi}(q)$ are both within C of $\mathscr{A}'$. It follows that $\mu \subset \mathscr{A}'$. Similarly $\mu \subset \mathscr{B}'$. Thus

$$\ell(\mathscr{A}' \cap \mathscr{B}') \geq \ell(\mu) \geq d_{\tilde{H}}(\tilde{\varphi}(p), \tilde{\varphi}(q)) - 2C.$$

Now if $x = b \cdot p \in \mathscr{A} \cap \mathscr{B}$, then the same reasoning gives

$$2C + \ell(\mathscr{A}' \cap \mathscr{B}') \geq d_{\tilde{H}}(\tilde{\varphi}(p), \tilde{\varphi}(x)) = d_{\tilde{H}}(\tilde{\varphi}(p), b \cdot \tilde{\varphi}(p)) \geq \ell(\beta|H).$$

Otherwise we have $d_{\tilde{G}}(q, x) \leq k$, so that

$$\begin{aligned}
\ell(\beta|H) \leq d_{\tilde{H}}(\tilde{\varphi}(p), b \cdot \tilde{\varphi}(p)) &\leq d_{\tilde{H}}(\tilde{\varphi}(p), \tilde{\varphi}(q)) + d_{\tilde{H}}(\tilde{\varphi}(q), \tilde{\varphi}(x)) \\
&\leq \left(\ell(\mathscr{A}' \cap \mathscr{B}') + 2C \right) + e^D k,
\end{aligned}$$

since $\tilde{\varphi}$ is e^D–Lipschitz. Thus β is $(2C + e^D k)$–almost contained in α at H. $\qquad\square$

We now make a simple observation. Recall from [Definition 6.4](#) that $\mathrm{right}_\gamma(\alpha)$ denotes the supremum of times along a folding path γ for which $\alpha|\gamma(t)$ contains an immersed illegal segment of length $\mathfrak{I}$.

Lemma 6.8. *Let* $\gamma \colon \mathbf{I} \to \mathfrak{X}$ *be a folding path and suppose that* $t \in \mathbf{I}$ *satisfies* $t \geq \mathrm{right}_\gamma(\alpha)$ *for some conjugacy class* α *of* $\mathbb{F}$. *If* β *is* k–*almost contained in* α *at* $G_t = \gamma(t)$ *and* $\ell(\beta|G_t) \geq 3k + 3\mathfrak{I}$, *then*

$$\mathrm{leg}(\beta|G_t) \geq \frac{2}{\mathfrak{I}} \ell(\beta|G_t).$$

Proof. The loop $\beta|G_t$ subdivides into two immersed subsegments β_0' and β_1', where β_0' has length at most k and β_1' is an immersed path into $\alpha|G_t$. By choosing the maximal length legal subsegments containing the two endpoints of these segments, we may alternately subdivide $\beta|G_t$ into 4 subsegments $\beta_0 p \beta_1 q$ separated at illegal turns, where β_1 is an immersed path into $\alpha|G_t$, β_0 has length at most k, and p and q are both legal. (Up to 3 of these segments may be degenerate, as happens in the case that $\beta|G_t$ is itself legal).

Since the endpoints of β_1 are at illegal turns, we may unambiguously talk about the legal length $\mathrm{leg}(\beta_1)$ of β_1 (defined in the same way as for conjugacy classes). Since β_1 is an immersed subpath in $\alpha|G_t$ and $t \geq \mathrm{right}_\gamma(\alpha)$ we see that every subpath of β_1 of length at least $\mathfrak{I}$ contains a legal segment of length at least 3. Therefore, if we subdivide β_1 into $n = \lfloor \ell(\beta_1)/\mathfrak{I} \rfloor$ subsegments of equal length $\ell(\beta_1)/n \geq \mathfrak{I}$, we see that each subsegment contains a legal segment of length 3. Therefore

$$\mathrm{leg}(\beta_1) \geq 3n \geq 3\left(\frac{\ell(\beta_1)}{\mathfrak{I}} - 1 \right) = \frac{3}{\mathfrak{I}} \ell(\beta_1) - 3.$$

Hence it follows that

$$\begin{aligned}
\mathrm{leg}(\beta|G_t) \geq \mathrm{leg}(\beta_1) + \ell(p) + \ell(q) &\geq \frac{3}{\mathfrak{I}} \ell(\beta_1) + \ell(p) + \ell(q) - 3 \\
&\geq \frac{3}{\mathfrak{I}}(\ell(\beta_1) + \ell(p) + \ell(q)) - 3 \geq \frac{3}{\mathfrak{I}}(\ell(\beta|G_t) - k) - 3.
\end{aligned}$$

Since $\ell(\beta|G_t) \geq 3k + 3\mathfrak{I}$ by hypothesis, we may conclude

$$\mathrm{leg}(\beta|G_t) \geq \frac{3}{\mathfrak{I}} \ell(\beta|G_t) - \left(\frac{3k}{\mathfrak{I}} + 3 \right) \geq \frac{3}{\mathfrak{I}} \ell(\beta|G_t) - \frac{1}{\mathfrak{I}} \ell(\beta|G_t) = \frac{2}{\mathfrak{I}} \ell(\beta|G_t). \qquad\square$$

Corollary 6.9. *Let $\gamma\colon \mathbf{I} \to \mathfrak{X}$ be a folding path, and suppose that β is k–almost contained in α at $G_s = \gamma(s)$. If $\ell(\beta) \geq 3k + 3\mathfrak{I}$ and $s \geq \mathrm{right}_\gamma(\alpha)$, then for all $t \in \mathbf{I}$ with $t \geq s$ we have*

$$\ell(\beta|G_t) \geq \frac{2}{3\mathfrak{I}}\ell(\beta|G_s)e^{t-s}.$$

Proof. By Lemma 6.2 and Lemma 6.8 we immediately see that

$$\ell(\beta|G_t) \geq \mathrm{leg}(\beta|G_t) \geq \frac{1}{3}\mathrm{leg}(\beta|G_s)e^{t-s} \geq \frac{2}{3\mathfrak{I}}\ell(\beta|G_s)e^{t-s}. \qquad \square$$

6.4 Flaring away from length minimizers

Given a subgroup $\Gamma \leq \mathrm{Out}(\mathbb{F})$, a point $R \in \mathfrak{X}$, and a conjugacy class α of $\mathbb{F}$, we write

$$\mathrm{m}_\Gamma^R(\alpha) := \inf\{\ell(\alpha|g\cdot R) : g \in \Gamma\}$$

for the infimal length of the conjugacy class α on the orbit $\Gamma \cdot R$. Observe that $\mathrm{m}_\Gamma^R(\alpha)$ is positive since it is bounded below by the injectivity radius of R. While this infimal length in principle need not be attained at any orbit point, we may nevertheless be assured that the set

$$\rho_\Gamma^R(\alpha) := \{g \in \Gamma : \ell(\alpha|g\cdot R) \leq 2\mathrm{m}_\Gamma^R(\alpha)\}$$

is nonempty.

We now come to the main technical lemma of §6, showing that if Γ is convex cocompact and β is long and almost contained in α at a point of $\rho_\Gamma^R(\alpha)$, then the length of β in the orbit $\Gamma \cdot R$ grows uniformly exponentially with the distance from $\rho_\Gamma^R(\alpha)$. In fact we show something slightly stronger than this:

Lemma 6.10. *Suppose that $\Gamma \leq \mathrm{Out}(\mathbb{F})$ is finitely generated with word metric d_Γ and that Γ qi-embeds into $\mathcal{F}$. For every $R \in \mathfrak{X}$ there exist constants $\lambda > 1$ and $C > 0$ such that for every $k > 0$ there is some $L_0 > 0$ with the following property: Let α be a simple conjugacy class of $\mathbb{F}$ and let $g_0 \in \Gamma$ lie in $\rho_\Gamma^R(\alpha)$. Suppose further that $g \in \Gamma$ lies on a geodesic from g_0 to $h \in \Gamma$, i.e. $d_\Gamma(g_0,h) = d_\Gamma(g_0,g) + d_\Gamma(g,h)$. If $\beta \in \mathbb{F}$ is k–almost contained in α at $g\cdot R$ and $\ell(\beta|g\cdot R) \geq L_0$, then*

$$\ell(\beta|h\cdot R) \geq C\lambda^{d_\Gamma(g,h)}\ell(\beta|g\cdot R).$$

Proof. Choose a free factor A in $\pi_{\mathcal{F}}(R)$. By assumption, the assignment $\varphi \mapsto \varphi \cdot A$ gives a quasi-isometric embedding $\Gamma \to \mathcal{F}$. Since $\pi_{\mathcal{F}}$ is coarsely Lipschitz (Lemma 2.5), it follows from Γ–equivariance that there is some K so that the assignment $\varphi \mapsto \varphi \cdot R$ gives a K–quasi-isometric embedding $\Gamma \to \mathfrak{X}$. Recall this means that

$$\frac{1}{K}d_\Gamma(\varphi,\psi) - K \leq d_\mathfrak{X}(\varphi\cdot R, \psi\cdot R) \leq Kd_\Gamma(\varphi,\psi) + K \tag{5}$$

for all $\varphi, \psi \in \Gamma$. Since $d_\Gamma(g_0,h) = d_\Gamma(g_0,g) + d_\Gamma(g,h)$, we may choose a geodesic $(\varphi_0,\ldots,\varphi_N)$ in (Γ, d_Γ) from $g_0 = \varphi_0$ to $h = \varphi_N$ with $g = \varphi_j$ for some $0 \leq j \leq N$. It then follows from (5) that the map $\gamma_0\colon [0,N] \to \mathfrak{X}$ defined by $\gamma_0(s) = \varphi_i \cdot R$ for $s \in [i,i+1) \cap [0,N]$ is a (directed) $3K$–quasigeodesic. As described in [BF2, Proposition 2.11] and [FM, Theorem 5.6], we may build a geodesic $\hat{\gamma}\colon \hat{\mathbf{I}} \to \mathfrak{X}$ from $g_0 \cdot R$ to $h \cdot R$ which is a concatenation of a *rescaling path* followed by a folding path. By Theorem 2.6, it follows that there exists constants $A, \varepsilon > 0$ and $K' \geq 1$ depending only on K and the injectivity radius of R such that $\gamma_0([0,N])$ and $\hat{\gamma}(\hat{\mathbf{I}})$ have Hausdorff distance at most A, that $\hat{\gamma}(\hat{\mathbf{I}}) \subset \mathfrak{X}_\varepsilon$, and that $\pi_{\mathcal{F}} \circ \hat{\gamma}\colon \hat{\mathbf{I}} \to \mathcal{F}$ is a K'–quasigeodesic. Furthermore, the rescaling portion of $\hat{\gamma}$ has length at most $\log(2/\varepsilon)$ by [DT, Lemma 2.6]. If $\gamma\colon \mathbf{I} \to \mathfrak{X}$ denotes the folding path portion of $\hat{\gamma}$, then after replacing A by $A + \mathsf{M}_\varepsilon \log(2/\varepsilon)$ (where M_ε is as in Lemma 2.4) we may conclude that $d_{\mathrm{Haus}}\big(\gamma(\mathbf{I}), \{\varphi_0\cdot R,\ldots,\varphi_N\cdot R\}\big) \leq A$, that $d_\mathfrak{X}^{\mathrm{sym}}(\gamma(\mathbf{I}_-), g_0\cdot R) \leq A$, that $\gamma(\mathbf{I}_+) = h\cdot R$, and that $\pi_{\mathcal{F}} \circ \gamma\colon \mathbf{I} \to \mathcal{F}$ is a K'–quasigeodesic. In particular, since $g = \varphi_j$, there is some $\mathbf{I}_0 \in \mathbf{I}$ for which $d_\mathfrak{X}^{\mathrm{sym}}(\gamma(\mathbf{I}_0), g\cdot R) \leq A$.

Conversely, for each $t \in \mathbf{I}$ there is some i so that $d_{\mathfrak{X}}^{\text{sym}}(\gamma(t), \varphi_i \cdot R) \leq A$ and consequently $\ell(\alpha|\varphi_i \cdot R) \leq e^A \ell(\alpha|\gamma(t))$. Similarly $\ell(\alpha|\gamma(\mathbf{I}_-)) \leq e^A \ell(\alpha|g_0 \cdot R)$. Furthermore, since $g_0 \in \rho_\Gamma^R(\alpha)$, the definition of $\rho_\Gamma^R(\alpha)$ gives $\ell(\alpha|g_0 \cdot R) \leq 2\ell(\alpha|\varphi_i \cdot R)$. Hence, for every $t \in \mathbf{I}$ we find that

$$\ell(\alpha|\gamma(\mathbf{I}_-)) \leq 2e^{2A}\ell(\alpha|\gamma(t)). \tag{6}$$

Let us analyze the location of $\text{right}_\gamma(\alpha)$ in $\mathbf{I}$. Set $D' = K'(\mathsf{B} + K')$, where B is the constant defined after Proposition 6.5. Then all $s,t \in \mathbf{I}$ with $|s-t| \geq D'$ satisfy $d_{\mathcal{F}}(\gamma(s), \gamma(t)) \geq \mathsf{B}$. We claim that

$$\text{left}_\gamma(\alpha) \leq \mathbf{I}_- + D'\left(\frac{2A}{\log 2} + 2\right).$$

Indeed, if this is false, then $\alpha|\gamma(t)$ is illegal for all $t \in [\mathbf{I}_-, \mathbf{I}_- + D'm]$, where $m = \left\lceil \frac{2A}{\log 2} + 1 \right\rceil \leq \frac{2A}{\log 2} + 2$. But then Lemma 6.3 would imply that $\ell(\alpha|\gamma(\mathbf{I}_-)) > 2^m \ell(\alpha|\gamma(\mathbf{I}_- + D'm))$, contradicting (6) since $2^m \geq 2e^{2A}$. We also know that $\text{right}_\gamma(\alpha) \leq \text{left}_\gamma(\alpha) + D'$ by Proposition 6.5 and our choice of D'. Therefore we conclude

$$\text{right}_\gamma(\alpha) \leq \mathbf{I}_- + D'\left(\frac{2A}{\log 2} + 3\right) \leq \mathbf{I}_0 + D'\left(\frac{2A}{\log 2} + 3\right). \tag{7}$$

Set $r_0 := \max\{\mathbf{I}_0, \text{right}_\gamma(\alpha)\}$, so that $\text{right}_\gamma(\alpha) \leq r_0 \leq \mathbf{I}_+$. Thus $d_{\mathfrak{X}}^{\text{sym}}(\gamma(\mathbf{I}_0), \gamma(r_0)) \leq \mathsf{M}_\varepsilon D'(\frac{2A}{\log 2} + 3)$. Combining this with $d_{\mathfrak{X}}^{\text{sym}}(\gamma(\mathbf{I}_0), g \cdot R) \leq A$ then gives $d_{\mathfrak{X}}^{\text{sym}}(\gamma(r_0), g \cdot R) \leq E$, where $E := \mathsf{M}_\varepsilon D'(\frac{2A}{\log 2} + 3) + A$.

Suppose now that β is k–almost contained in α at $g \cdot R$ for the given constant k. By Lemma 6.7, we know that β is k'–almost contained α at $\gamma(r_0)$ for some constant k' depending only on k and E. Define $L_0 := (3k' + 3\mathfrak{I})e^E$ so that the additional assumption $\ell(\beta|g \cdot R) \geq L_0$ will moreover imply

$$\ell(\beta|\gamma(r_0)) \geq \ell(\beta|g \cdot R)e^{-E} \geq 3k' + 3\mathfrak{I}.$$

Thus if β is k–almost contained in α at $g \cdot R$ and $\ell(\beta|g \cdot R) \geq L_0$, we may apply Corollary 6.9 to conclude

$$\ell(\beta|\gamma(t)) \geq \frac{2}{3\mathfrak{I}}\ell(\beta|\gamma(r_0))e^{(t-r_0)} \geq \frac{2}{3\mathfrak{I}e^E}\ell(\beta|g \cdot R)e^{(t-r_0)} \tag{8}$$

for all $t \geq r_0$ (this is valid because $r_0 \geq \text{right}_\gamma(\alpha)$).

We now use (8) to prove the proposition. Since $\mathbf{I}_+ \geq r_0$ by construction, (8) immediately gives

$$\ell(\beta|h \cdot R) = \ell(\beta|\gamma(\mathbf{I}_+)) \geq \frac{2}{3\mathfrak{I}e^E}\ell(\beta|g \cdot R)e^{(\mathbf{I}_+ - r_0)}.$$

Since $d_{\mathfrak{X}}^{\text{sym}}(g \cdot R, \gamma(r_0)) \leq E$ and γ is a directed geodesic, we have that $d_{\mathfrak{X}}(g \cdot R, h \cdot R) \leq E + (\mathbf{I}_+ - r_0)$. We also know $d_{\mathfrak{X}}(g \cdot R, h \cdot R) \geq \frac{1}{K}d_\Gamma(g,h) - K$ by Equation (5). Therefore $(\mathbf{I}_+ - r_0) \geq \frac{1}{K}d_\Gamma(g,h) - K - E$ and so

$$\ell(\beta|h \cdot R) \geq \frac{2}{3\mathfrak{I}e^{2E+K}}\ell(\beta|g \cdot R)\left(e^{1/K}\right)^{d_\Gamma(g,h)}$$

as desired. Thus the conclusion of the proposition holds with $\lambda = e^{1/K}$ and $C = \frac{2}{3\mathfrak{I}e^{2E+K}}$. □

7 Distortion within fibers of $E_\Gamma \to \Gamma$

Fix a finitely generated subgroup $\Gamma \leq \text{Out}(\mathbb{F})$ for which the orbit map $\Gamma \to \mathcal{CS}$ is a quasi-isometric embedding. Then by Theorem 1.1 and Theorem 4.12, the corresponding extension E_Γ is hyperbolic. In this section, we establish Theorem 7.9 which shows that if $H \leq \mathbb{F}$ is finitely generated and of infinite index, then H is quasiconvex (and hence undistorted) as a subgroup of E_Γ. This will follow from the structural result Theorem 7.2, which will be used in §8 to characterize which hyperbolic extensions of $\mathbb{F}$ induce convex cocompact subgroups of $\text{Out}(\mathbb{F})$.

7.1 The Cayley graph bundle

To this end, we first recall some notation and results from [DT, §§7–8] describing the structure of E_Γ. Fix a finite generating set $S = \{s_1, \ldots, s_n\}$ of Γ and a free basis $X = \{x_1, \ldots, x_r\}$ for $\mathbb{F}$. Recalling that E_Γ is the preimage of Γ under the quotient map $\mathrm{Aut}(\mathbb{F}) \to \mathrm{Out}(\mathbb{F})$, we choose a lift $t_i \in \mathrm{Aut}(\mathbb{F})$ of s_i for each $1 \leq i \leq n$. In general, we will use the notation $\tilde{g} \in \mathrm{Aut}(\mathbb{F})$ to denote a lift of $g \in \Gamma$ to E_Γ. We also write $i_x \in \mathrm{Aut}(\mathbb{F})$ for the inner automorphism given by conjugation by $x \in \mathbb{F}$, i.e., $i_x(a) = xax^{-1}$ for $a \in \mathbb{F}$. Note that $\varphi i_x \varphi^{-1} = i_{\varphi(x)}$ for each $x \in \mathbb{F}$ and $\varphi \in \mathrm{Aut}(\mathbb{F})$.

As a subgroup of $\mathrm{Aut}(\mathbb{F})$, E_Γ is thus generated by the set $W = \{i_{x_1}, \ldots i_{x_r}, t_1, \ldots, t_n\}$. That is

$$E_\Gamma = \langle i_{x_1}, \ldots i_{x_r}, t_1, \ldots, t_n \rangle \leq \mathrm{Aut}(\mathbb{F}).$$

For convenience, set $\hat{X} = \{i_{x_1}, \ldots, i_{x_r}\}$ and $\hat{\mathbb{F}} = \langle \hat{X} \rangle$, so that $\hat{\mathbb{F}}$ is the isomorphic image of $\mathbb{F}$ in $\mathrm{Aut}(\mathbb{F})$. Note that $\hat{\mathbb{F}}$ is also the kernel of the homomorphism $E_\Gamma \to \Gamma$. We also set $\tilde{S} = \{t_1, \ldots, t_n\}$.

Let $T = \mathrm{Cay}(\mathbb{F}, X)$, $\mathscr{E} = \mathrm{Cay}(E_\Gamma, W)$, and $\mathscr{B} = \mathrm{Cay}(\Gamma, S)$, where $\mathrm{Cay}(\cdot, \cdot)$ denotes the Cayley graph with the specified generating set equipped with the path metric in which each edge has length one. We respectively view $\hat{\mathbb{F}} \cong \mathbb{F}$, E_Γ, and Γ as the 0–skeletons of the simplicial complexes T, $\mathscr{E}$, and $\mathscr{B}$. Set $\mathcal{R}$ to be the standard rose on the generating set X so that $\mathcal{R} = T/\mathbb{F}$. There is then an obvious equivariant simplicial map

$$p \colon \mathscr{E} \to \mathscr{B}$$

extending the surjective homomorphism $E_\Gamma \to \Gamma$; note that p sends edges of $\mathscr{E}$ to either vertices or edges of $\mathscr{B}$ depending on whether the edge corresponds to a generator in $\hat{X}$ or $\tilde{S}$, respectively. For each $b \in \Gamma$, we see that the preimage $T_b = p^{-1}(b)$ is the simplicial tree (isomorphic to T) with vertices labeled by the coset $\tilde{b}\hat{\mathbb{F}}$ ($\tilde{b}$ any lift of b) and edges labeled by $\hat{X}$. We write d_b for the induced path metric on the fiber T_b over $b \in \Gamma$. By a k–qi section of $p \colon \mathscr{E} \to \mathscr{B}$, we simply mean a k–quasi-isometric embedding $\sigma \colon \mathscr{B} \to \mathscr{E}$ such that $(p \circ \sigma)(g) = g$ for every $g \in \Gamma$ (i.e., for every vertex of $\mathscr{B}$). By Mosher's "quasi-isometric section lemma" [Mos], there exists a constant $\mathfrak{K} \geq 1$ (depending only on the bundle $p \colon \mathscr{E} \to \mathscr{B}$) such that for every $b \in \Gamma$ and vertex $v \in T_b$ one may build a $\mathfrak{K}$–qi section $\sigma \colon \mathscr{B} \to \mathscr{E}$ with $\sigma(b) = v$.

As discussed in [DT, §7] (see also [MS, Example 1.8]), $p \colon \mathscr{E} \to \mathscr{B}$ is an example of the *metric graph bundle* construction developed by Mj and Sardar in [MS]. In particular, there is a **metric properness function** $f \colon \mathbb{N} \to \mathbb{N}$ such that $d_b(u, v) \leq f(d_{\mathscr{E}}(u, v))$ for all $b \in \Gamma$ and all vertices $u, v \in T_b$ [DT, Lemma 7.2]. We moreover observe that if group elements $u, v \in E_\Gamma$ lie in the same fiber T_b, then $u^{-1}v \in \hat{\mathbb{F}}$ and the fiberwise distance satisfies $d_b(u, v) = \left| u^{-1}v \right|_{\hat{X}}$ [DT, Lemma 7.1]. Writing $u^{-1}v = i_a$ for the appropriate $a \in \mathbb{F}$ and defining $\mathfrak{L} \geq 1$ to be the maximal bilipschitz constant of the automorphisms in $\tilde{S}$, we deduce the inequality

$$d_{bs_i}(ut_i, vt_i) = \left| t_i^{-1}u^{-1}vt_i \right|_{\hat{X}} = \left| t_i^{-1}i_a t_i \right|_{\hat{X}} = \left| i_{t_i^{-1}(a)} \right|_{\hat{X}} \leq \mathfrak{L} |i_a|_{\hat{X}} = \mathfrak{L}d_b(u, v) \tag{9}$$

for every generator $s_i \in S$ with corresponding lift $t_i \in \tilde{S}$.

7.2 The Width Theorem

Suppose that a is an *element* (rather than a conjugacy class) of $\mathbb{F}$. Then left multiplication by the inner automorphism $i_a \in \hat{\mathbb{F}}$ gives an isometry of $\mathscr{E}$ that preserves each fiber T_b of $p \colon \mathscr{E} \to \mathscr{B}$. In particular, i_a acts as a hyperbolic isometry of (T_b, d_b) translating along a unique invariant axis. We write $\mathscr{A}_b(a)$ for the axis of i_a in the fiber T_b and then define the **axis bundle of** a to be the union $\mathscr{A}(a) := \cup_{b \in \Gamma} \mathscr{A}_b(a)$.

Note that while $\mathscr{A}_g(a)$ is a geodesic in the path metric space (T_g, d_g), it will generally be far from being a geodesic in the whole space $\mathscr{E}$. However, our next result shows that when a is simple and $g \in \Gamma$ lies in the minimizing set $\rho_\Gamma^R([a])$ for the conjugacy class of a (see §6.4), then $\mathscr{A}_g(a)$ is a uniform quasigeodesic in $\mathscr{E}$.

Proposition 7.1. *Suppose that $\Gamma \leq \mathrm{Out}(\mathbb{F})$ qi-embeds into $\mathcal{CS}$ and let $p \colon \mathscr{E} \to \mathscr{B}$ be as in §7.1. Then for every $R \in \mathfrak{X}$ there exists a constant $Q \geq 1$ such that for any simple element $a \in \mathbb{F}$ and any $g \in \rho_\Gamma^R([a])$, the axis $\mathscr{A}_g(a)$ (viewed as a map $\mathbb{R} \to T_g \subset \mathscr{E}$) is a Q–quasigeodesic in $\mathscr{E}$.*

The proof of Proposition 7.1 is fairly technical and will be deferred to the next section. Meanwhile, we use it to uniformly bound the "width" of all simple conjugacy classes of $\mathbb{F}$. Suppose now that $\Gamma \leq \mathrm{Out}(\mathbb{F})$ qi-embeds into $\mathbb{CS}$ so that the corresponding bundle $\mathscr{E}$ is a hyperbolic metric space. Every element $a \in \mathbb{F}$ then acts (via left multiplication by $i_a \in \hat{\mathbb{F}}$) as a hyperbolic isometry of $\mathscr{E}$, and we let a^* denote a biinfinite geodesic of $\mathscr{E}$ joining the two fixed points $a^\infty, a^{-\infty}$ of a in $\partial \mathscr{E}$. Define the **width** of $a \in \mathbb{F}$ (or its conjugacy class α) by

$$\mathrm{width}(\alpha) = \mathrm{width}(a) := \mathrm{diam}_{\mathscr{B}} \, p(a^*).$$

Theorem 7.2 (Width Theorem). *Suppose that $\Gamma \leq \mathrm{Out}(\mathbb{F})$ qi-embeds into $\mathbb{CS}$ and consider the hyperbolic extension $p\colon E_\Gamma \to \Gamma$ of $\mathbb{F}$. Then the simple conjugacy classes of $\mathbb{F}$ have uniformly bounded width. That is,*

$$\sup_\alpha \mathrm{diam}_\Gamma \, p(\alpha^*) < \infty$$

where the supremum is over simple conjugacy classes of $\mathbb{F}$.

Proof. Let a be a simple element of $\mathbb{F}$ and α is conjugacy class. Suppose that the length of α is minimized over the fiber T_g for $g \in \Gamma$. Then by Proposition 7.1, the axis $\mathscr{A}_g(a)$ of a in T_g is a Q–quasigeodesic for $Q \geq 0$ not depending on a. As E_Γ is hyperbolic, Proposition 2.1 provides a constant $R = R(Q) \geq 0$ so that the axis of a in E_Γ has Hausdorff distance at most R from any Q–quasigeodesic joining its endpoints in ∂E_Γ. Hence, $d_{\mathrm{Haus}}(a^*, \mathscr{A}_g(a)) \leq R$ and so the diameter of the image of a^* in Γ is at most R. Since this is independent of the conjugacy class α, the theorem follows. $\qquad \square$

In §8, we will show that any hyperbolic extension of $\mathbb{F}$ in which simple elements have uniformly bounded width is an extension by a convex cocompact subgroup $\Gamma \leq \mathrm{Out}(\mathbb{F})$.

7.3 Axis bundles and the proof of Proposition 7.1

We now embark on the proof of Proposition 7.1. Our approach is modeled on that of Kent–Leininger in [KL4], where they prove an analogous result in order to establish their width theorem for hyperbolic extensions of surface groups. The main idea is to use the axis bundle $\mathscr{A}(a)$ to construct a Lipschitz retract from $\mathscr{E}$ to $\mathscr{A}_g(a)$.

The first step of the construction utilizes the techniques that Mitra [Mit2] has developed and used extensively to study hyperbolic group extensions. Let $p\colon \mathscr{E} \to \mathscr{B}$ be as in §7.1 and let $a \in \mathbb{F}$ be a nontrivial element with corresponding axis bundle $\mathscr{A}(a) \subset \mathscr{E}$. Define $\mathfrak{p}_a\colon E_\Gamma \to \mathscr{A}(a)$ to be the fiber-wise closest point projection to $\mathscr{A}(a)$, that is, for each vertex x in the fiber T_g we define $\mathfrak{p}_a(x)$ to be the unique point in $\mathscr{A}_g(a)$ minimizing the distance $d_g(x, \mathfrak{p}_a(x))$.

Lemma 7.3 (Mitra [Mit2, Lemma 3.2]). *There is a constant $\mathfrak{C} \geq 1$ depending only on the bundle $\mathscr{E} \to \mathscr{B}$ such that $\mathfrak{p}_a\colon E_\Gamma \to \mathscr{A}(a)$ is $\mathfrak{C}$–Lipschitz for each element $a \in \mathbb{F}$. That is, for for all $u, v \in E_\Gamma$ we have*

$$d_{\mathscr{E}}(\mathfrak{p}_a(u), \mathfrak{p}_a(v)) \leq \mathfrak{C} d_{\mathscr{E}}(u, v).$$

Proof. Mitra's proof of Lemma 7.3 follows from basic hyperbolic geometry; for completeness we give a brief sketch here: By the triangle inequality it suffices to suppose $d_{\mathscr{E}}(u, v) = 1$. Then if u, v lie in the same fiber T_g, we immediately have $d_{\mathscr{E}}(\mathfrak{p}_a(u), \mathfrak{p}_a(v)) \leq d_g(\mathfrak{p}_a(u), \mathfrak{p}_a(v)) \leq 1$ by the nature of closest-point-projection in the tree T_g. Otherwise u and v lie in neighboring fibers so that $v = ut_i$ for some $t_i \in \tilde{S}$. But then one may use the uniform bilipschitz equivalence of neighboring fibers (Equation (9)) to prove that $\mathfrak{p}_a(u)t_i$ is uniformly close to $\mathfrak{p}_a(v)$ (which is the content of [Mit2, Lemma 3.6]). $\qquad \square$

Lemma 7.3 allows us to extend $\mathfrak{p}_a$ to a map $\mathfrak{p}_a\colon \mathscr{E} \to \mathscr{A}(a)$ that is coarsely $\mathfrak{C}$–Lipschitz. For $a \in \mathbb{F}$ nontrivial, let us use the terminology k–**qi section through** $\mathscr{A}(a)$ to mean a k–quasi-isometric embedding $\sigma\colon (\mathscr{B}, d_{\mathscr{B}}) \to (\mathscr{E}, d_{\mathscr{E}})$ such that $\sigma(g) \in \mathscr{A}_g(a)$ for all $g \in \Gamma$.

Corollary 7.4. *For any nontrivial $a \in \mathbb{F}$ and any vertex $v \in \mathscr{A}_g(a)$, there exists a $\mathfrak{C}\mathfrak{K}$–qi section σ through $\mathscr{A}(a)$ with $\sigma(g) = v$.*

Proof. Let $\sigma_0 \colon \mathscr{B} \to \mathscr{E}$ be the $\mathfrak{K}$–qi section with $\sigma_0(g) = v$ provided by Mosher [Mos]. Composing σ_0 with $\mathfrak{p}_a \colon \mathscr{E} \to \mathscr{A}(a)$ then gives the desired $\mathfrak{C}\mathfrak{K}$–quasi-isometric embedding $\mathfrak{p}_a \circ \sigma_0 \colon \mathscr{B} \to \mathscr{E}$. $\square$

We now make a basic observation about "well-separated" k–qi sections through axis bundles.

Lemma 7.5. *There exists a constant $\mathfrak{D} > 0$ depending only on $\mathscr{E} \to \mathscr{B}$ with the following property. Suppose $a \in \mathbb{F}$ is nontrivial, that σ_1, σ_2 are $\mathfrak{C}\mathfrak{K}$–qi sections through $\mathscr{A}(a)$, and that $g, h \in \Gamma$ satisfy $d_{\mathscr{B}}(g, h) \le 1$. If $u \in \mathscr{A}_g(a)$ lies between $\sigma_1(g)$ and $\sigma_2(g)$ on $\mathscr{A}_g(a)$ with $d_g(u, \sigma_i(g)) \ge \mathfrak{D}$ for $i = 1, 2$ and $v \in \mathscr{A}_h(a)$ is a vertex with $d_{\mathscr{E}}(u, v) \le \mathfrak{C}\mathfrak{K}$, then v also lies between $\sigma_1(h)$ and $\sigma_2(h)$ on $\mathscr{A}_h(a)$.*

Proof. Define $\mathfrak{D} := 4E^2$, where $E = \mathfrak{L} + 2R_0(\mathfrak{L}, 0) + f(2\mathfrak{C}\mathfrak{K} + 1 + R_0(\mathfrak{L}, 0))$, and suppose that $g, h, u, v, \sigma_1, \sigma_2$ are as in the statement of the lemma. Here R_0 is as in Proposition 2.1. If $d_{\mathscr{B}}(g, h) = 0$ then v is a point on the geodesic $\mathscr{A}_g(a)$ within d_g–distance $f(\mathfrak{C}\mathfrak{K})$ of u. Thus the result is immediate since $\mathfrak{D} > f(\mathfrak{C}\mathfrak{K})$. Otherwise $d_{\mathscr{B}}(g, h) = 1$ so that $h = gs_i$ for some generator $s_i \in S$. Let t_i be the chosen lift in $\tilde{S}$. Define a map Ψ from $\tilde{g}\hat{\mathbb{F}}$ (the vertex set of T_g) to $\tilde{h}\hat{\mathbb{F}}$ (the vertex set of T_h) by declaring $\Psi(x)$ to be the d_h–closest-point-projection of $xt_i \in T_h$ to $\mathscr{A}_h(a)$. Since the assignment $x \mapsto xt_i$ is $\mathfrak{L}$–bilipschitz by Equation (9) and T_h is 0–hyperbolic, it follows from Proposition 2.1 and the definition of $E > \mathfrak{L} + 2R_0(\mathfrak{L}, 0)$ that Ψ restricts to an E–quasi-isometric embedding from (the vertices of) $\mathscr{A}_g(a)$ to $\mathscr{A}_h(a)$. Observe also that $d_{\mathscr{E}}(\Psi(x), x) \le 1 + R_0(\mathfrak{L}, 0)$ for each vertex $x \in \mathscr{A}_g(a)$.

The hypotheses on $\sigma_1(g), u, \sigma_2(g)$ now imply that $\Psi(u)$ appears between $\Psi(\sigma_1(g))$ and $\Psi(\sigma_2(g))$ on $\mathscr{A}_h(a)$ with $d_h(\Psi(u), \Psi(\sigma_i(g))) \ge 3E$ for $i = 1, 2$. Using the triangle inequality, the hypotheses on σ_1, σ_2, v with the above observation about $\Psi|_{\mathscr{A}_g(a)}$ together give

$$d_{\mathscr{E}}\big(\Psi(\sigma_1(g)), \sigma_1(h)\big),\, d_{\mathscr{E}}\big(\Psi(\sigma_2(g)), \sigma_2(h)\big),\, d_{\mathscr{E}}\big(\Psi(u), v\big) \le 2\mathfrak{C}\mathfrak{K} + 1 + R_0(\mathfrak{L}, 0).$$

In particular, the first inequality follows from the estimate

$$\begin{aligned}
d_{\mathscr{E}}\big(\Psi(\sigma_1(g)), \sigma_1(h)\big) &\le d_{\mathscr{E}}\big(\Psi(\sigma_1(g)), \sigma_1(g)\big) + d_{\mathscr{E}}\big(\sigma_1(g), \sigma_1(h)\big) \\
&\le 1 + R_0(\mathfrak{L}, 0) + \mathfrak{C}\mathfrak{K} \cdot d_{\mathscr{B}}(g, h) + \mathfrak{C}\mathfrak{K} \\
&\le 2\mathfrak{C}\mathfrak{K} + 1 + R_0(\mathfrak{L}, 0),
\end{aligned}$$

and the other two are similar. By metric properness, we may thus conclude that

$$d_h\big(\Psi(\sigma_1(g)), \sigma_1(h)\big),\, d_h\big(\Psi(\sigma_2(g)), \sigma_2(h)\big),\, d_h\big(\Psi(u), v\big) \le f(2\mathfrak{C}\mathfrak{K} + 1 + R_0(\mathfrak{L}, 0)) \le E.$$

Therefore the triangle inequality shows that v lies between $\sigma_1(h)$ and $\sigma_2(h)$ on $\mathscr{A}_h(a)$, as desired. $\square$

Next, when $a \in \mathbb{F}$ is simple the flaring property established in §6.4 translates into the following estimate for well-separated qi-sections through $\mathscr{A}(a)$. Let $\mathscr{R} \in \mathcal{X}$ denotes the marked graph $\mathrm{Cay}(\mathbb{F}, X)/\mathbb{F}$ equipped with the metric in which each edge has length $1/\mathrm{rk}(\mathbb{F})$

Proposition 7.6. *Suppose that $\Gamma \le \mathrm{Out}(\mathbb{F})$ qi-embeds into $\mathcal{F}$ and let $p \colon \mathscr{E} \to \mathscr{B}$ be as in §7.1. For every $\bar{K} \ge 1$ there exist $D, D_0 > 0$ and $\eta > 1$ so that the following holds. Suppose that $a \in \mathbb{F}$ is simple and that σ, σ' are $\bar{K}$–qi-sections through $\mathscr{A}(a)$ with $d_g(\sigma(g), \sigma'(g)) \ge D_0$ for some element $g \in \rho_\Gamma^{\mathscr{R}}([a])$. Then for all $h \in \Gamma$ we have:*

$$d_h(\sigma(h), \sigma'(h)) \ge D\eta^{d_\Gamma(g, h)} d_g(\sigma(g), \sigma'(g)).$$

Proof. Le C, λ, and L_0 be the constants obtained by applying Lemma 6.10 to the orbit $\Gamma \cdot \mathscr{R}$ with $k = 1$, and let f be the metric properness function for the graph bundle $\mathscr{E} \to \mathscr{B}$. Fix $N \ge 1$ large enough so that $C\lambda^N > 4$ and define

$$D_0 := \max\left\{2,\, 2\,\mathrm{rk}(\mathbb{F})L_0,\, 8C^{-1}f(6\bar{K}N)\right\}.$$

Choose $h \in \Gamma$ arbitrarily, and let $g = g_0, \ldots, g_m = h$ be a geodesic from g to h in Γ. Let us write $\sigma_i := \sigma(g_i) \in T_{g_i}$ and $\sigma_i' := \sigma'(g_i) \in T_{g_i}$ for the value of the two $\bar{K}$–qi-sections in the fiber T_{g_i} of $p \colon \mathscr{E} \to \mathscr{B}$ over g_i. Choose any vertices g_i, g_j along our geodesic with $i \le j \le i + 2N$, and suppose temporarily that $d_{g_i}(\sigma_i, \sigma_i') \ge D_0$. Recall from §7.1 that $T_{g_i} = p^{-1}(g_i)$ is a simplicial tree whose edges are labeled by the free basis $\hat{X}$ of $\hat{\mathbb{F}}$. With respect to this basis, the element $i_c = \sigma_i^{-1}\sigma_i' \in \hat{\mathbb{F}}$ may not by cyclically reduced. However, there is some $x \in \hat{X}$ so that $i_b = \sigma_i^{-1}\sigma_i'x \in \hat{\mathbb{F}}$ is cyclically reduced. Set $z_i = \sigma_i$ and $z_i' = \sigma_i'x$, so that the geodesic edge path $[z_i, z_i']$ in T_{g_i} is labeled by the cyclically reduced word i_b of $\hat{\mathbb{F}}$ with the properties that $\|i_b\|_{\hat{X}} = |i_b|_{\hat{X}}$ and that $|i_b|_{\hat{X}}$ differs from $d_{g_i}(\sigma_i, \sigma_i') = \left|\sigma_i^{-1}, \sigma_i'\right|_{\hat{X}}$ by at most 1.

Choosing a lift $\tilde{g}_i \in \mathrm{Aut}(\mathbb{F})$ of $g_i \in \mathrm{Out}(\mathbb{F})$, the action $\tilde{g}_i$ on $\mathscr{E}$ restricts to a simplicial automorphism from T_1 to T_{g_i} that respects the edge labeling and thus gives an identification of $T_1 = \mathrm{Cay}(\hat{\mathbb{F}}, \hat{X})$ with T_{g_i}. With respect to this identification, the element $\tilde{g}_i(b) \in \mathbb{F}$ acts on T_{g_i} the same way that i_b acts on $\mathrm{Cay}(\hat{\mathbb{F}}, \hat{X})$. Therefore, since the labeled edge path $[z_i, z_i']$ is a fundamental domain of the axis of i_b in $\mathrm{Cay}(\hat{\mathbb{F}}, \hat{X})$, it follows that $[z_i, z_i'] \subset T_{g_i}$ is a fundamental domain of the axis for $\tilde{g}_i(b)$ acting on T_{g_i}. Letting β and α denote the conjugacy classes of b and a, respectively, it follows that $\tilde{g}_i(\beta) = g_i(\beta)$ is 1–almost contained in α at $g_i \cdot \mathcal{R}$ (since $\sigma_i, \sigma_i' \in \mathscr{A}_{g_i}(a)$ by construction and all edges in the universal cover of $g_i \cdot \mathcal{R}$ have length $1/\mathrm{rk}(\mathbb{F}) \le 1$). Since

$$\ell(g_i(\beta)|g_i \cdot \mathcal{R}) = \ell(\beta|\mathcal{R}) = \tfrac{1}{\mathrm{rk}(\mathbb{F})}\|i_b\|_{\hat{X}} \ge \tfrac{1}{\mathrm{rk}(\mathbb{F})}\left(d_{g_i}(\sigma_i, \sigma_i') - 1\right) \ge \tfrac{1}{2\mathrm{rk}(\mathbb{F})}d_{g_i}(\sigma_i, \sigma_i') \ge L_0$$

by the assumption $d_{g_i}(\sigma_i, \sigma_i') \ge D_0$, we may apply Lemma 6.10 to conclude

$$\ell(g_i(\beta)|g_j \cdot \mathcal{R}) \ge C\lambda^{j-i}\ell(g_i(\beta)|g_i \cdot \mathcal{R}) \ge \tfrac{C}{2\mathrm{rk}(\mathbb{F})}\lambda^{j-i}d_{g_i}(\sigma_i, \sigma_i').$$

For each $i < p \le j$, set $s_p = g_{p-1}^{-1}g_p \in S$ and let $t_p \in \tilde{S}$ be the chosen lift of s_p in the generating set W of E_Γ. For $i \le p \le j$ let us also define $z_p = z_i t_{i+1}\cdots t_p$ and $z_p' = z_i' t_{i+1}\cdots t_p$, both of which are points over $g_p = g_i s_{i+1}\cdots s_p$. Since $g_i, g_{i+1}, \ldots, g_j$ is a geodesic in Γ, it follows that $z_i, \ldots, z_j$ and $z_i', \ldots, z_j'$ are both geodesics in E_Γ and thus that $d_{\mathscr{E}}(z_i, z_j) = d_{\mathscr{E}}(z_i', z_j') = j - i \le 2N$. Observe now that

$$z_j^{-1}z_j' = (z_i t_{i+1}\cdots t_j)^{-1}\left(z_i' t_{i+1}\cdots t_j\right) = \left(t_j^{-1}\cdots t_{i+1}^{-1}\right)z_i^{-1}z_i'(t_{i+1}\cdots t_j) = \varphi i_b \varphi^{-1} = i_{\varphi(b)} \in \hat{\mathbb{F}},$$

where φ is the specific lift $\varphi = t_j^{-1}\cdots t_{i+1}^{-1} \in \mathrm{Aut}(\mathbb{F})$ of $g_j^{-1}g_i = s_j^{-1}\cdots s_{i+1}^{-1}$. In particular, we see that the distance between z_j and z_j' in the fiber T_{g_j} satisfies

$$d_{g_j}(z_j, z_j') = \left|i_{\varphi(b)}\right|_{\hat{X}} \ge \left\|i_{\varphi(b)}\right\|_{\hat{X}} = \mathrm{rk}(\mathbb{F})\ell(\varphi(\beta)|\mathcal{R}) = \mathrm{rk}(\mathbb{F})\ell(g_i(\beta)|g_j \cdot \mathcal{R}) \ge \tfrac{C}{2}\lambda^{j-i}d_{g_i}(\sigma_i, \sigma_i'). \quad (10)$$

Let us now compare this distance to $d_{g_j}(\sigma_j, \sigma_j')$. Since σ, σ' are $\bar{K}$–qi-sections, the quantities $d_{\mathscr{E}}(\sigma_i, \sigma_j)$ and $d_{\mathscr{E}}(\sigma_i', \sigma_j')$ are bounded by $\bar{K}(j - i) + \bar{K} \le 2\bar{K}N + \bar{K}$. Since we also have $d_{\mathscr{E}}(\sigma_i, z_i), d_{\mathscr{E}}(\sigma_i', z_i') \le 1$ by construction, the triangle inequality thus gives

$$d_{\mathscr{E}}(z_j, \sigma_j) \le d_{\mathscr{E}}(z_j, z_i) + d_{\mathscr{E}}(z_i, \sigma_i) + d_{\mathscr{E}}(\sigma_i, \sigma_j)$$
$$\le 2N + 1 + 2\bar{K}N + \bar{K} \le 6\bar{K}N,$$

and similarly, $d_{\mathscr{E}}(z_j', \sigma_j') \le 6\bar{K}N$. By metric properness of the graph bundle $\mathscr{E} \to \mathscr{B}$, it follows that $d_{g_j}(z_j, \sigma_j), d_{g_j}(z_j', \sigma_j') \le f(6\bar{K}N)$. Combining with (10) and using $d_{g_i}(\sigma_i, \sigma_i') \ge D_0 \ge 8C^{-1}f(6\bar{K}N)$ and $\lambda^{j-i} \ge 1$, we conclude that

$$d_{g_j}(\sigma_j, \sigma_j') \ge d_{g_j}(z_j, z_j') - d_{g_j}(z_j, \sigma_j) - d_{g_j}(z_j', \sigma_j')$$
$$\ge \tfrac{C}{2}\lambda^{j-i}d_{g_i}(\sigma_i, \sigma_i') - 2f(6\bar{K}N)$$
$$\ge \left(\lambda^{j-i} - \tfrac{1}{2}\right)\tfrac{C}{2}d_{g_i}(\sigma_i, \sigma_i') \ge \tfrac{C}{4}\lambda^{j-i}d_{g_i}(\sigma_i, \sigma_i').$$

To summarize, we have now shown that the implication

$$d_{g_i}(\sigma_i, \sigma_i') \geq D_0 \implies d_{g_j}(\sigma_j, \sigma_j') \geq \tfrac{C}{4} \lambda^{j-i} d_{g_i}(\sigma_i, \sigma_i') \tag{11}$$

holds for any pair of vertices g_i, g_j on the geodesic $g = g_0, \ldots, g_m = h$ with $i \leq j \leq i + 2N$. Suppose now that $d_g(\sigma(g), \sigma'(g)) \geq D_0$. If $d_\Gamma(g, h) > N$, we may then break the geodesic $g_0, \ldots, g_m$ into $\lfloor d_\Gamma(g,h)/N \rfloor \geq \frac{1}{2N} d_\Gamma(g,h)$ pieces that each have length between N and $2N$ and inductively apply the estimate (11) to conclude

$$d_h(\sigma(h), \sigma'(h)) \geq \left(\tfrac{C}{4} \lambda^N\right)^{\lfloor d_\Gamma(g,h)/N \rfloor} d_g(\sigma(g), \sigma'(g)) \geq \eta^{d_\Gamma(g,h)} d_g(\sigma(g), \sigma'(g)),$$

where $\eta := \left(\frac{C\lambda^N}{4}\right)^{1/2N} > 1$. Otherwise, $d_\Gamma(g, h) \leq N$ and (11) immediately gives the desired bound

$$d_h(\sigma(h), \sigma'(h)) \geq \left(\tfrac{C}{4\eta^N}\right) \eta^{d_\Gamma(g,h)} d_g(\sigma(g), \sigma'(g)).$$

Setting $D = \frac{C}{4\eta^N}$ completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

With these tools in hand, we are now prepared to give the

Proof of Proposition 7.1. Let $\mathfrak{K}$, $\mathfrak{C}$, and $\mathfrak{D}$ be the constants provided by Mosher's quasi-isometric section lemma, Lemma 7.3, and Lemma 7.5 (all which depend only on the bundle $\mathscr{E} \to \mathscr{B}$). Let $D, D_0 > 0$ be the constants obtained by applying Proposition 7.6 with $\bar{K} = \mathfrak{C}\mathfrak{K}$, and fix a constant $M > D_0 + (f(\mathfrak{C}) + \mathfrak{D})/D$.

Let $a \in \mathbb{F}$ and $g \in \rho_\Gamma^{\mathcal{R}}([a])$ be as in the statement of the proposition. We may use Corollary 7.4 to construct an infinite family $\{\Sigma_i\}_{i \in \mathbb{Z}}$ of $\mathfrak{C}\mathfrak{K}$–qi sections through $\mathscr{A}(a)$ with the property that and $d_g(\Sigma_i(g), \Sigma_j(g)) = M|i - j|$ for all $i, j \in \mathbb{Z}$. Notice that this forces the points $\ldots, \Sigma_{-1}(g), \Sigma_0(g), \Sigma_1(g), \ldots$ to be linearly ordered along the axis $\mathscr{A}_g(a)$. Furthermore, for all $i \neq j$ we have $d_g(\Sigma_i(g), \Sigma_j(g)) \geq D_0$ so that we may apply Proposition 7.6 to conclude $d_h(\Sigma_i(h), \Sigma_j(h)) \geq DM > \mathfrak{D} + f(\mathfrak{C})$ for all $h \in \Gamma$.

From this we claim that the sections $\{\Sigma_i\}$ are *consistently ordered* in each fiber, meaning that if $\Sigma_j(h)$ appears between $\Sigma_i(h)$ and $\Sigma_k(h)$ in the axis $\mathscr{A}_h(a)$ for some $h \in \Gamma$, then the same holds for every $h' \in \Gamma$. Indeed, if $\Sigma_j(h)$ appears between $\Sigma_i(h)$ and $\Sigma_k(h)$ in $\mathscr{A}_h(a)$, then applying Lemma 7.5 with $\sigma_1 = \Sigma_i$, $\sigma_2 = \Sigma_k$, $u = \Sigma_j(h)$ shows that $v = \Sigma_j(h')$ appears between $\Sigma_i(h')$ and $\Sigma_k(h)$ in any neighboring fiber h'; thus the consistently ordered conclusion follows by induction.

We now use the sections $\{\Sigma_i\}_{i \in \mathbb{Z}}$ to define a map $\mathfrak{q}_a \colon \mathscr{A}(a) \to \mathscr{A}_g(a)$, as follows. For each $h \in \Gamma$, the sections Σ_i partition the geodesic $\mathscr{A}_h(a) \cong \mathbb{R}$ into infinitely many, disjoint, half open intervals $Z_j^h := [\Sigma_j(h), \Sigma_{j+1}(h))$. Define the map $\mathfrak{q}_a$ by sending the interval $Z_j^h = [\Sigma_j(h), \Sigma_{j+1}(h))$ to the point $\Sigma_j(g) \in \mathscr{A}_g(a)$ (so the image of $\mathfrak{q}_a$ is the set $\{\Sigma_j(g)\}_{j \in \mathbb{Z}}$). Next define $\Pi_a \colon E_\Gamma \to \mathscr{A}_g(a)$ to be the composition $\Pi_a = \mathfrak{q}_a \circ \mathfrak{p}_a$.

Claim. *The map $\Pi_a \colon E_\Gamma \to (\mathscr{A}_g(a), d_g)$ is a coarse $3M$–Lipschitz retraction onto $\mathscr{A}_g(a)$, meaning that for all $u, v \in E_\Gamma$ and each vertex $x \in \mathscr{A}_g(a)$ we have*

$$d_g(\Pi_a(u), \Pi_a(v)) \leq 3M d_{\mathscr{E}}(u, v) \qquad and \qquad d_g(x, \Pi_a(x)) \leq 3M.$$

Indeed, for any vertex $x \in \mathscr{A}_g(a)$ we have $\mathfrak{p}_a(x) = x$, by definition, so that $\Pi_a(x) = \Sigma_j(g)$ where $j \in \mathbb{Z}$ is the unique integer such that $x \in Z_j^g = [\Sigma_j(g), \Sigma_{j+1}(g))$. Thus $d_g(x, \Pi_a(x)) \leq M$ since $d_g(\Sigma_j(g), \Sigma_{j+1}(g)) = M$ by construction of the family $\{\Sigma_i\}$. To complete the proof of the claim, it thus suffices to prove the bound $d_g(\Pi_a(u), \Pi_a(v)) \leq 3M$ for all $u, v \in E_\Gamma$ with $d_{\mathscr{E}}(u, v) = 1$. First suppose $u, v \in T_h$ for some $h \in \Gamma$. Then $\mathfrak{p}_a(u), \mathfrak{p}_a(v) \in \mathscr{A}_h(a)$ satisfy

$$d_h(\mathfrak{p}_a(u), \mathfrak{p}_a(v)) \leq f\big(d_{\mathscr{E}}(\mathfrak{p}_a(u), \mathfrak{p}_a(v))\big) \leq f(\mathfrak{C})$$

by Lemma 7.3. Since $d_h(\Sigma_m(h), \Sigma_n(h)) \geq \mathfrak{D} + f(\mathfrak{C})$ for all $m \neq n$, it follows that if $\mathfrak{p}_a(u)$ lies in Z_i^h and $\mathfrak{p}_a(v)$ lies in Z_j^h, then $|i - j| \leq 1$. Thus $d_g(\Pi_a(u), \Pi_a(v)) = d_g(\Sigma_i(g), \Sigma_j(g)) = M|i - j| \leq M$. Next suppose that u and v lie in different fibers. Then, since $d_{\mathscr{E}}(u, v) = 1$, we have $u \in T_h$ and $v \in T_{h'}$ for some $h, h' \in \Gamma$

with $d_{\mathscr{B}}(h,h') = 1$. Let $i,j \in \mathbb{Z}$ be such that $\mathfrak{p}_a(u) \in Z_i^h$ and $\mathfrak{p}_a(v) \in Z_j^{h'}$, and note that $d_{\mathscr{E}}(\mathfrak{p}_a(u),\mathfrak{p}_a(v)) \leq \mathfrak{C}$ by Lemma 7.3. Since $\mathfrak{p}_a(u) \in [\Sigma_i(h),\Sigma_{i+1}(h))$, it follows that $\mathfrak{p}_a(u)$ lies between $\Sigma_{i-1}(h)$ and $\Sigma_{i+2}(h)$ with $d_h(\mathfrak{p}_a(u),\Sigma_n(h)) \geq \mathfrak{D}$ for $n \in \{i-1,i+2\}$. Therefore we may apply Lemma 7.5 to conclude that $\mathfrak{p}_a(v)$ lies between $\Sigma_{i-1}(h')$ and $\Sigma_{i+2}(h')$ in $\mathscr{A}_{h'}(a)$. In particular, we must have $j \in \{i-1,i,i+1,i+2\}$ so that $d_g(\Pi_a(u),\Pi_a(v)) = d_g(\Sigma_i(g),\Sigma_j(g)) = M|i-j| \leq 3M$. This completes the proof of the claim.

We now prove the proposition. Let $x,y \in \mathscr{A}_g(a)$ be arbitrary. Then clearly $d_{\mathscr{E}}(x,y) \leq d_g(x,y)$ by definition of the path metrics $d_{\mathscr{E}}$ and d_g. Choosing vertices $x',y' \in \mathscr{A}_a(g)$ with $d_g(x,x'),d_g(y,y') \leq 1$, the claim and triangle inequality together imply that

$$d_g(x,y) \leq 2 + 6M + d_g(\Pi_a(x'),\Pi_a(y')) \leq 2 + 6M + 3M d_{\mathscr{E}}(x',y') \leq 2 + 6M + 3M(d_{\mathscr{E}}(x,y)+2).$$

Therefore the inclusion $(\mathscr{A}_g(a),d_g) \to (\mathscr{E},d_{\mathscr{E}})$ is a $(12M+2)$–quasi-isometric embedding. $\qquad\square$

7.4   A Scott–Swarup theorem

In [SS], Scott and Swarup proved that a finitely generated, infinite index subgroup of the fiber of a fibered hyperbolic 3-manifold group is quasiconvex. This result was extended to arbitrary hyperbolic extensions of surface groups in [DKL] and to hyperbolic free-by-cyclic groups with fully irreducible monodromy in [Mit3]. In this section, we generalize these results on the nondistortion of finitely generated, infinite index subgroups of fiber group to the case of hyperbolic extensions of free group by convex cocompact subgroups of $\mathrm{Out}(\mathbb{F})$.

We first show each free factor A of $\mathbb{F}$ is undistorted in E_Γ. Our proof uses the following well-known fact about hyperbolic groups:

Fact 7.7. *Suppose that G is a hyperbolic group and let $a,b \in G$ be infinite order elements. Then $(a^n b^{-n})^\infty \to a^\infty$ and $(a^n b^{-n})^{-\infty} \to b^\infty$ in ∂G as $n \to \infty$.*

Proposition 7.8. *Let $\Gamma \leq \mathrm{Out}(\mathbb{F})$ be a finitely generated group with quasi-isometric orbit map into CS and let E_Γ be the associated hyperbolic extensions of $\mathbb{F}$. Then for any free factor A of $\mathbb{F}$, A is quasiconvex in E_Γ.*

Proof. Let δ be the hyperbolicity constant of E_Γ. By Proposition 7.1 and Theorem 7.2, there are constants $R,Q > 1$ so that $\mathrm{diam}_\Gamma(p(x^*)) \leq R$ for each simple element x of $\mathbb{F}$ and that x^* and the Q–quasigeodesic $\mathscr{A}_g(x)$ have Hausdorff distance at most R whenever x is minimized in the fiber over g, i.e. whenever $g \in \rho_\Gamma^{\mathscr{R}}([x])$.

Now let $a,b \in A$ be arbitrary. We claim that $d_\Gamma(g,h) \leq 5R + 4\delta$ for any $g \in \rho_\Gamma^{\mathscr{R}}([a])$ and $h \in \rho_\Gamma^{\mathscr{R}}([b])$. Since $a^n b^{-n} \in A$, these elements are simple. Moreover, since $(a^n b^{-n})^\infty \to a^\infty$ and $(a^n b^{-n})^{-\infty} \to b^\infty$ in ∂E_Γ by Fact 7.7, there is an $N \geq 0$ such that $(a^N b^{-N})^*$ meets a 2δ–neighborhood of a^* and b^* in E_Γ. Then

$$\mathrm{diam}(p(a^*) \cup p(b^*)) \leq \mathrm{diam}\, p(a^*) + 2\delta + \mathrm{diam}\, p((a^N b^{-N})^*) + 2\delta + \mathrm{diam}\, p(b^*) \leq 3R + 4\delta. \qquad (12)$$

Therefore $d_\Gamma(g,h) \leq 5R + 4\delta$ as claimed. Setting $D = d_\Gamma(1,g) + 5R + 4\delta$, this moreover shows that $d_\Gamma(1,h) \leq D$ whenever h lies in the minimizing set $\rho_\Gamma^{\mathscr{R}}([b])$ for any $b \in A$.

We can now directly verify that A is quasiconvex in E_Γ. Identify A with the vertices of the tree T_1^A in $\mathscr{E}$. For any two vertices a,b of T_1^A there is an $x \in A$ whose axis $\mathscr{A}_1(x)$ in T_1 passes through the vertices a and b. If x is minimized in the fiber over $h \in \Gamma$, then $d_\Gamma(1,h) \leq D$. By Equation (9), there exists an $\mathbb{F}$–equivariant $\mathscr{L}^D$–bilipschitz map $T_1 \to T_h$ (obtained by writing h as a geodesic $s_1 \cdots s_n$ in Γ and lifting the $s_i \in S$ to generators $t_i \in \tilde{S}$). This and the fact that $\mathscr{A}_h(x)$ is a Q–quasigeodesic together imply that $\mathscr{A}_1(x)$ is a Q'–quasigeodesic in $\mathscr{E}$ for some constant $Q' \geq 1$ depending only on Q, $\mathscr{L}$, and D. Hence, by Proposition 2.1, any geodesic joining a and b in $E_\Gamma = \mathscr{E}^0$ stays within the $R_0(Q',\delta)$–neighborhood of the geodesic $\mathscr{A}_1(x)$ in T_1^A joining a and b. Therefore A is quasiconvex in E_Γ. $\qquad\square$

We can now combine Proposition 7.8 with Proposition 5.2 to prove the main result of this section.

Theorem 7.9 (Nondistortion in fibers)**.** *Suppose that $\Gamma \leq \mathrm{Out}(\mathbb{F})$ quasi-isometrically embeds into CS, and let L be a finitely generated subgroup of the fiber $\mathbb{F} \lhd E_\Gamma$. Then L is quasiconvex, and hence undistorted, in the hyperbolic extension E_Γ if and only if L has infinite index in $\mathbb{F}$.*

Proof. Suppose that L is a finitely generated, infinite index subgroup of $\mathbb{F}$. By Marshall Hall's theorem, L is a free factor of H for some finite index subgroup $H \leq \mathbb{F}$. By Proposition 5.2, the group Γ^H qi-embeds into $\mathcal{CS}(H)$, and hence the corresponding H–extension E_{Γ^H} fitting into the sequence

$$1 \longrightarrow H \longrightarrow E_{\Gamma^H} \longrightarrow \Gamma^H \longrightarrow 1$$

is hyperbolic by Theorem 1.1 and Theorem 4.12. Since L is a free factor of H, Proposition 7.8 implies that L is quasiconvex in E_{Γ^H}. Finally, since H has finite index in $\mathbb{F}$, E_Γ and E_{Γ^H} are commensurable, and we conclude that L is quasiconvex in E_Γ. Conversely, if L has finite index in $\mathbb{F}$, then L is quasi-isometric to $\mathbb{F} \lhd E_\Gamma$ which itself is exponentially distorted in E_Γ by virtue of being infinite and normal. $\qquad\square$

Remark 7.10. We note that the above theorem does not necessarily hold for hyperbolic extensions of $\mathbb{F}$ by groups that do not admit quasi-isometric orbit maps into $\mathcal{CS}$. For example, if ϕ is an automorphism of $\mathbb{F}$ which is atoroidal but fixes the conjugacy class of a free factor A, then the $\mathbb{F}$–extension $\mathbb{F} \rtimes \langle \phi \rangle$ is hyperbolic by [Bri], but the subgroup A is not quasiconvex.

8 Hyperbolicity of E_Γ and convex cocompactness of Γ

In the previous section, we learned that if Γ is convex cocompact and purely atoroidal then not only is the extension E_Γ hyperbolic, but the projection $E_\Gamma \to \Gamma$ has controlled geometry over the axes of simple elements. In this section, we develop a converse to the main theorem of [DT], which established hyperbolicity of E_Γ. That is, we impose additional structural properties of E_Γ that imply the induced orbit map $\Gamma \to \mathcal{F}$ is a quasi-isometric embedding. These properties turn out to characterize convex cocompact subgroups of $\mathrm{Out}(\mathbb{F})$ among the class of subgroups inducing hyperbolic extensions of $\mathbb{F}$.

Suppose henceforth that $1 \to \mathbb{F} \to E \xrightarrow{p} Q \to 1$ is a hyperbolic extension of $\mathbb{F}$. This short exact sequence induces an outer action of Q on $\mathbb{F}$ given by the homomorphism $Q \to \mathrm{Out}(\mathbb{F})$ sending $q \in Q$ to the class of the automorphism that conjugates $\mathbb{F} \lhd E$ by any lift $\tilde{q} \in E$ of q. We then have the commutative diagram

$$
\begin{array}{ccccccccc}
1 & \longrightarrow & \mathbb{F} & \longrightarrow & E & \xrightarrow{p} & Q & \longrightarrow & 1 \\
& & \| & & \downarrow & & \downarrow & & \\
1 & \longrightarrow & \mathbb{F} & \longrightarrow & E_\Gamma & \longrightarrow & \Gamma & \longrightarrow & 1,
\end{array}
\tag{13}
$$

where Γ is the image of $Q \to \mathrm{Out}(\mathbb{F})$. Fixing finite generating sets for E and Q, for each element $a \in \mathbb{F}$ we continue to write a^* for a geodesic in E joining $a^{-\infty} \in \partial E$ to $a^\infty \in \partial E$. The image $p(a^*)$ in Q then depends only on the $\mathbb{F}$–conjugacy class α of a. Hence, as in §7.2, we may define the **width** of $a \in \mathbb{F}$ (or α) to be

$$\mathrm{width}_Q(a) = \mathrm{width}_Q(\alpha) := \mathrm{diam}_Q\, p(\alpha^*).$$

Theorem 8.1 (Convex cocompactness). *Suppose that $1 \to \mathbb{F} \to E \to Q \to 1$ is a hyperbolic extension of $\mathbb{F}$. Then Q has convex cocompact image in $\mathrm{Out}(\mathbb{F})$ (and hence admits a quasi-isometric embedding orbit map into $\mathcal{CS}$) if and only there exists $D \geq 0$ so that $\mathrm{width}_Q(a) \leq D$ for each simple element $a \in \mathbb{F}$.*

Proof. Since E is hyperbolic, the induced homomorphism $\chi \colon Q \to \mathrm{Out}(\mathbb{F})$ must have finite kernel, otherwise E contains a subgroup isomorphic to $\mathbb{F} \times \ker(\chi)$ (see [DT, §2.5]). Thus Q is quasi-isometric to its image Γ in $\mathrm{Out}(\mathbb{F})$ and, further, each vertical arrow in (13) has finite kernel. From this we see that E is $\mathbb{F}$–equivariantly quasi-isometric to E_Γ and moreover that $\mathrm{width}_Q(a)$ coarsely agrees with the width $\mathrm{width}(a)$ in Γ as defined in §7.2. Therefore if Γ is convex cocompact, Theorem 7.2 shows that supremum $\sup_a \mathrm{width}_Q(a)$ over all simple $a \in \mathbb{F}$ is bounded.

For the converse, suppose that $\mathrm{width}_Q(\beta) \leq D$ for each simple conjugacy class β of $\mathbb{F}$. Since E_Γ is δ–hyperbolic, Γ is purely atoroidal and so it suffices to show that Γ, or equivalently Q, qi-embeds into $\mathcal{F}$ by Theorem 4.12. As it is more natural for our argument, we instead work with the quasi-isometric primitive

loop graph $\mathcal{PL}$ defined in §4. Fix $\alpha \in \mathcal{PL}^0$ and consider the orbit map $Q \to \mathcal{PL}$ given by $g \mapsto g \cdot \alpha$, where Q acts on $\mathcal{PL}$ via $Q \to \mathrm{Out}(\mathbb{F})$. We define a coarse map $\sigma \colon \mathcal{PL} \to Q$ which we show is a coarse Lipschitz retraction for the orbit map $Q \to \mathcal{PL}$. Since the orbit map is necessarily Lipschitz, this will show that $Q \to \mathcal{PL}$ is a quasi-isometric embedding and establish the theorem. Set

$$\sigma(\beta) = p(\beta^*),$$

which is by assumption a subset of Q of diameter at most D. This map is equivariant since for each $g \in Q$ and any lift $\tilde{g} \in E$,

$$\sigma(g \cdot \beta) = \sigma(\tilde{g}b) = p(\tilde{g}b^*) = g p(\beta^*),$$

where b is any representative of the conjugacy class of β. Hence, if we set $D_0 = \mathrm{diam}_Q(\{1\} \cup \sigma(\alpha))$, then $\sigma \colon \mathcal{PL} \to \Gamma$ is indeed a D_0–coarse retraction and so it only remains to show that it is Lipschitz.

Let β and γ be adjacent conjugacy classes in $\mathcal{PL}$ and choose representatives b and c such that $\langle b, c \rangle$ is a rank 2 free factor of $\mathbb{F}$. Then, for each $n \in \mathbb{Z}$, $b^n c^n$ is a simple element of $\mathbb{F}$, and by Fact 7.7 $(b^n c^n)^\infty$ approaches b^∞ as $n \to \infty$ and $(b^n c^n)^{-\infty}$ approaches $c^{-\infty}$ as $n \to \infty$. Hence, the axis $(b^n c^n)^*$ in E becomes forward asymptotic to b^* and backward asymptotic to c^*. Then, just as in Equation (12), for all sufficiently large n we have that

$$\mathrm{diam}_Q(p(b^*) \cup p(c^*)) \leq \mathrm{diam}_Q\, p(b^*) + 2\delta + \mathrm{diam}_Q\, p((b^n c^n)^*) + 2\delta + \mathrm{diam}_Q\, p(c^*) \leq 3D + 4\delta.$$

This demonstrates that $\sigma \colon \mathcal{PL} \to Q$ is a Lipschitz retract and completes the proof. $\square$

References

[AKB]   Yael Algom-Kfir and Mladen Bestvina. Asymmetry of outer space. *Geom. Dedicata*, 156(1):81–92, 2012.

[BF1]   Mladen Bestvina and Mark Feighn. Outer limits. Preprint (http://andromeda.rutgers.edu/~feighn/papers/outer.pdf), 1994.

[BF2]   Mladen Bestvina and Mark Feighn. Hyperbolicity of the complex of free factors. *Adv. Math.*, 256:104–155, 2014.

[BFH]   Mladen Bestvina, Mark Feighn, and Michael Handel. Laminations, trees, and irreducible automorphisms of free groups. *Geom. Funct. Anal.*, 7(2):215–244, 1997.

[BH1]   Mladen Bestvina and Michael Handel. Train tracks and automorphisms of free groups. *Ann. of Math. (2)*, 135(1):1–51, 1992.

[BH2]   Martin R. Bridson and Aandre Haefliger. *Metric spaces of non-positive curvature*, volume 319. Springer, 2009.

[BR]   Mladen Bestvina and Patrick Reynolds. The boundary of the complex of free factors. *Duke Math. J.*, 164(11):2213–2251, 2015.

[Bri]   Peter Brinkmann. Hyperbolic automorphisms of free groups. *Geom. Funct. Anal.*, 10(5):1071–1089, 2000.

[CHL]   Thierry Coulbois, Arnaud Hilion, and Martin Lustig. $\mathbb{R}$-trees and laminations for free groups. II. The dual lamination of an $\mathbb{R}$-tree. *J. Lond. Math. Soc. (2)*, 78(3):737–754, 2008.

[CHR]   Thierry Coulbois, Arnaud Hilion, and Patrick Reynolds. Indecomposable F_N-trees and minimal laminations. *Groups Geom. Dyn.*, 9(2):567–597, 2015.

[CL]   Marshall M. Cohen and Martin Lustig. Very small group actions on **R**-trees and Dehn twist auto-morphisms. *Topology*, 34(3):575–617, 1995.

[CV]   Marc Culler and Karen Vogtmann. Moduli of graphs and automorphisms of free groups. *Invent. Math.*, 84(1):91–119, 1986.

[DKL]   Spencer Dowdall, Richard P. Kent, IV, and Christopher J Leininger. Pseudo-anosov subgroups of fibered 3-manifold groups. *Groups, Geometry, and Dynamics*, 8(4):1247–1282, 2014.

[DKT]   Spencer Dowdall, Ilya Kapovich, and Samuel J. Taylor. Cannon–thurston maps for hyperbolic free group extensions. *Israel J. Math.*, 216(2):753–797, 2016.

[DT]   Spencer Dowdall and Samuel J Taylor. Hyperbolic extensions of free groups. To appear in *Geom. Topol.* Preprint arXiv:1406.2567, 2014.

[FM]   Stefano Francaviglia and Armando Martino. Metric properties of outer space. *Publ. Mat.*, 55(2):433–473, 2011.

[GdlH]   É. Ghys and P. de la Harpe, editors. *Sur les groupes hyperboliques d'après Mikhael Gromov*, volume 83 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1990. Papers from the Swiss Seminar on Hyperbolic Groups held in Bern, 1988.

[GJLL]   Damien Gaboriau, Andre Jaeger, Gilbert Levitt, and Martin Lustig. An index for counting fixed points of automorphisms of free groups. *Duke Math. J.*, 93(3):425–452, 1998.

[Gui]   Vincent Guirardel. Actions of finitely generated groups on $\mathbb{R}$-trees. *Ann. Inst. Fourier (Grenoble)*, 58(1):159–211, 2008.

[Ham1]   Ursula Hamenstädt. Word hyperbolic extensions of surface groups. arXiv:math/0505244, 2005.

[Ham2]   Ursula Hamenstädt. Stability of quasi-geodesics in Teichmüller space. *Geom. Dedicata*, 146:101–116, 2010.

[Ham3]   Ursula Hamenstädt. The boundary of the free factor graph and the free splitting graph. Preprint arXiv:1211.1630, 2013.

[HH]   Ursula Hamenstädt and Sebastian Hensel. Convex cocompact subgroups of $\mathrm{Out}(F_n)$. Preprint arXiv:1411.2281, 2014.

[HM]   Michael Handel and Lee Mosher. The expansion factors of an outer automorphism and its inverse. *Trans. Amer. Math. Soc.*, 359(7):3185–3208, 2007.

[Kap1]   Ilya Kapovich. Currents on free groups. In *Topological and asymptotic aspects of group theory*, volume 394 of *Contemp. Math.*, pages 149–176. Amer. Math. Soc., Providence, RI, 2006.

[Kap2]   Ilya Kapovich. On purely loxodromic actions. *Monatshefte für Mathematik*, pages 1–13, 2015.

[KB]   Ilya Kapovich and Nadia Benakli. Boundaries of hyperbolic groups. In *Combinatorial and geometric group theory (New York, 2000/Hoboken, NJ, 2001)*, volume 296 of *Contemp. Math.*, pages 39–93. Amer. Math. Soc., Providence, RI, 2002.

[KL1]   Ilya Kapovich and Martin Lustig. Geometric intersection number and analogues of the curve complex for free groups. *Geom. Topol.*, 13(3):1805–1833, 2009.

[KL2]   Ilya Kapovich and Martin Lustig. Intersection form, laminations and currents on free groups. *Geom. Funct. Anal.*, 19(5):1426–1467, 2010.

[KL3]   Richard P. Kent, IV and Christopher J. Leininger. Shadows of mapping class groups: capturing convex cocompactness. *Geom. Funct. Anal.*, 18(4):1270–1325, 2008.

[KL4]   Richard P. Kent, IV and Christopher J. Leininger. A geometric criterion to be pseudo-Anosov. *Michigan Math. J.*, 63(2):227–251, 2014.

[Kob]   Tsuyoshi Kobayashi. Heights of simple loops and pseudo-anosov homeomorphisms. In *Braids*, volume 78, pages 327–338. Amer. Math. Soc., Providence, RI, Contemp. Math., 1988.

[KR]   Ilya Kapovich and Kasra Rafi. On the hyperbolicity of the free splitting and free factor complexes. *Groups Geom. Dyn.*, 8(2):391–414, 2014.

[Man]   Brian Mann. *Some hyperbolic Out(F_N)-graphs and nonunique ergodicity of very small F_N-trees*. ProQuest LLC, Ann Arbor, MI, 2014. Thesis (Ph.D.)–The University of Utah.

[Mit1] Mahan Mitra. Ending laminations for hyperbolic group extensions. *Geom. Funct. Anal.*, 7(2):379–402, 1997.

[Mit2] Mahan Mitra. Cannon-Thurston maps for hyperbolic group extensions. *Topology*, 37(3):527–538, 1998.

[Mit3] Mahan Mitra. On a theorem of Scott and Swarup. *Proc. Amer. Math. Soc*, 127(6):1625–1631, 1999.

[MM]   Howard A. Masur and Yair N. Minsky. Geometry of the complex of curves. I. Hyperbolicity. *Invent. Math.*, 138(1):103–149, 1999.

[Mos]   Lee Mosher. Hyperbolic extensions of groups. *J. Pure Appl. Algebra*, 110(3):305–314, 1996.

[MR1]   Brian Mann and Patrick Reynolds. *In preparation*.

[MR2]   Mahan Mj and Kasra Rafi. Algebraic ending laminations and quasiconvexity. Preprint arXiv:1506.08036v2, 2015.

[MS]   Mahan Mj and Pranab Sardar. A combination theorem for metric bundles. *Geom. Funct. Anal.*, 22(6):1636–1707, 2012.

[Rey1] Patrick Reynolds. On indecomposable trees in the boundary of outer space. *Geometriae Dedicata*, 153(1):59–71, 2011.

[Rey2] Patrick Reynolds. Reducing systems for very small trees. Preprint arXiv:1211.3378, 2012.

[SS]   Peter G. Scott and Gadde A. Swarup. Geometric finiteness of certain Kleinian groups. *Proc. Amer. Math. Soc*, 109(3):765–768, 1990.

Department of Mathematics   Department of Mathematics
Vanderbilt University   Yale University
1326 Stevenson Center   10 Hillhouse Ave
Nashville, TN 37240 , U.S.A   New Haven, CT 06520, U.S.A
E-mail: `spencer.dowdall@vanderbilt.edu`   E-mail: `s.taylor@yale.edu`