Constructing stable images

Jean Pierre Mutanguha*

January 9, 2022

Abstract

There is an algorithm for constructing a canonical representative for an injective free group endomorphism. The main corollary to our algorithm is an affirmative answer to Ventura's question: yes, the stable image for a free group endomorphism can be computed. This corollary also generalizes to all finite rank free groups a result due to Ciobanu–Logan in rank 2. By work of Bogopolski–Maslakova, it implies that the fixed point subgroup of a free group endomorphism can be computed. The final corollary is that the hyperbolicity of an ascending HNN extension of a free group can be algorithmically determined by looking solely at the dynamics of the defining monodromy.

Introduction

Let $\phi\colon F \to F$ be an endomorphism of a finitely generated free group. The stable image of ϕ, denoted $\phi^\infty(F)$, is the intersection of all the iterated images $\phi^i(F)$ for $i \geq 1$. This image has rank bounded above by the rank of F. In fact, Turner [Tur96] showed that it is a retract of F and, if ϕ is injective, a free factor of F. By the Hopfian property of finitely generated free groups, the restriction of ϕ to the stable image is an automorphism. Consequently, the stable image can be used to reduce questions about a free group endomorphism to questions about a free group automorphism. This is how Imrich–Turner [IT89] extended Bestvina–Handel's proof [BH92] of *Scott's conjecture* to free group endomorphisms. Scott's conjecture — now Bestvina–Handel's theorem — states that if ϕ is an automorphism, then the fixed point subgroup $\mathrm{Fix}(\phi) = \{x \in F : \phi(x) = x\}$ has rank bounded above by the rank of F.

On the computational side, Bogopolski–Maslakova gave an algorithm that computes a basis for $\mathrm{Fix}(\phi)$ when ϕ is an automorphism [BM16]; later, Feighn–Handel gave another algorithm [FH18, Proposition 9.10]. In the general case where ϕ is just an endomorphism, being able to algorithmically compute a basis for $\phi^\infty(F)$ would combine with either of

* *Email:* mutanguha@mpim-bonn.mpg.de, *Web address:* https://mutanguha.com
Max Planck Institute for Mathematics, Bonn, Germany

these algorithms to give one that computes a basis for $\mathrm{Fix}(\phi)$. Ciobanu–Logan recently gave an algorithm that computes bases for $\mathrm{Fix}(\phi)$ and $\phi^\infty(F)$ for any endomorphism ϕ of a rank 2 free group [CL22]; however, the higher rank cases remained open.

In previous work [Mut21], we studied the dynamics of injective free group *outer* endomorphisms. We proved that an injective outer endomorphism $[\phi]$ has a canonical representative $f\colon (\Gamma, \mathcal{G}) \to (\Gamma, \mathcal{G})$ on a *free splitting* $(\Gamma, \mathcal{G})$ of F with the following useful properties:

- the f-periodic vertex groups in $\mathcal{G}$ correspond to a $[\phi]$-fixed free factor system of F.
- f lifts to a ϕ-equivariant expanding immersion on the Bass–Serre tree for $(\Gamma, \mathcal{G})$.

It follows immediately that the fixed system from the first property is the unique maximal $[\phi]$-fixed free factor system. Free splittings will be defined in the next section but the term used will be simply *graphs*, short for *graphs of roses*. Canonical representatives will also be referred to as *automorphic expansions* where *automorphic* means the first property and *expansion* means the second. When $[\phi]$ is an outer automorphism, then the $(\Gamma, \mathcal{G})$ is the trivial free splitting consisting of a singleton labelled by F. So these representatives are only interesting when $[\phi]$ is not an outer automorphism.

We originally needed such a representative to prove a hyperbolization theorem for the *mapping torus* of ϕ, also known as an *ascending HNN extension* of F. However, the existence proof was not algorithmic since such considerations were irrelevant to hyperbolization. The main result of the present paper is giving an effective proof for this existence theorem:

Theorem (Constructing canonical representatives)**.** *There is an algorithm that takes an injective free group endomorphism as input and outputs its canonical representative.*

Let us, for a brief moment, discuss what made the previous proof nonconstructive. That proof consisted of two main steps: first, we proved the existence of a unique maximal fixed free factor system; then we used the uniqueness and maximality of the system to construct the canonical representative. Compartmentalizing the proof like this made the proof conceptually easier to follow. But here is the (algorithmic) issue: the way to certify that we have the maximal fixed free factor system is by exhibiting the canonical representative; but we cannot construct the representative using this proof without first being sure we have the right system. This becomes a chicken-and-egg problem!

The way around it is to construct the maximal fixed free factor system and canonical representative simultaneously: construct a fixed free factor system, use it to construct a piece of the canonical representative, use this partial representative to extend the fixed free factor system, use the larger fixed free factor system to extend the partial representative... The back-and-forth ends when we have the complete representative. The cost to this approach is the proof might be conceptually harder to follow. This summary does not even address how to construct a fixed free factor system in the first place. We will sketch the effective proof a little more carefully at the end of this introduction.

On the other hand, the effective proof is actually more elementary as it makes no use of Bestvina–Handel's train track theory, only Stallings folds and bounded cancellation. A second somewhat subtle improvement is that we construct a representive for the endomorphism ϕ, rather than the outer endomorphism $[\phi]$: the free splitting $(\Gamma, \mathcal{G})$ has a basepoint fixed by the representative f. By forgetting the basepoint and restricting f to the *core* of $(\Gamma, \mathcal{G})$, we would get the canonical representative for $[\phi]$. We do not prove uniqueness of this representative for ϕ as it is irrelevant to our applications; however, the proof is essentially the same as that for $[\phi]$ in [Mut21, Proposition 4.6].

Returning to the question of computing stable images and assuming ϕ is injective. Let $f\colon (\Gamma, \mathcal{G}) \to (\Gamma, \mathcal{G})$ be the canonical representative for ϕ. It follows from the definitions that the vertex group labelling the basepoint of $(\Gamma, \mathcal{G})$ corresponds to the stable image $\phi^\infty(F)$. In particular, our effective proof immediately gives us a way to compute stable images of injective free group endomorphisms. It is a consequence of the Hopfian property that this gives us a way to compute stable images of free group endomorphisms. Applying Bogopolski–Maslakova's algorithm to stable images allows us to compute fixed point subgroups as well; this answers an old open question, see [Ven10, Problem 1].

Corollary (Constructing stable images)**.** *There is an algorithm that takes a free group endomorphism as input and outputs a basis for its stable image.*

Corollary (Constructing fixed point subgroups)**.** *There is an algorithm that takes a free group endomorphism as input and outputs a basis for its fixed point subgroup.*

The last corollary of our effective proof is that the hyperbolicity of ϕ's mapping torus can be determined by studying only ϕ's dynamics.

Corollary (Detecting hyperbolicity)**.** *There is an algorithm that takes an injective free group endomorphism as input and (correctly) determines whether its mapping torus is word-hyperbolic.*

In fact, the last corollary applies more generally to HNN extensions of free groups over free factors. We will now end the introduction with the careful sketch as promised.

Sketch of effective proof. Assume ϕ is not surjective and proceed by induction on the *complexity* of F. We use nonsurjectivity and bounded cancellation to find a $[\phi]$-invariant proper free factor system $\mathcal{G}_1$, a free splitting $(\Gamma_1, \mathcal{G}_1)$, and a *relative immersion* $f_1\colon (\Gamma_1, \mathcal{G}_1) \to (\Gamma_1, \mathcal{G}_1)$, i.e. a representative that lifts to a ϕ-equivariant immersion on the Bass–Serre tree (Theorem 3.1, descend). Since $\mathcal{G}_1$ is a proper free factor system, the restriction of ϕ to $\mathcal{G}_1$ has a canonical representative $f_2\colon (\Gamma_2, \mathcal{G}_2) \to (\Gamma_2, \mathcal{G}_2)$ by the induction hypothesis. Blow-up the vertices of $(\Gamma_1, \mathcal{G}_1)$ by replacing them with the corresponding components of $(\Gamma_2, \mathcal{G}_2)$. This gives us a free splitting $(\Gamma', \mathcal{G}_2)$ with $(\Gamma_2, \mathcal{G}_2)$ as a subgraph and a representative $f'\colon (\Gamma', \mathcal{G}_2) \to (\Gamma', \mathcal{G}_2)$ whose restriction to $(\Gamma_2, \mathcal{G}_2)$ is f_2 and such that collapsing $(\Gamma_2, \mathcal{G}_2)$ recovers f_1.

Using the fact that f_2 is a canonical representative and f_1 a relative immersion, we replace $(\Gamma', \mathcal{G}_2)$ with another free splitting that has the same vertex groups and assume f' is a relative immersion whose periodic vertex groups in $\mathcal{G}_2$ correspond to a proper $[\phi]$-fixed free factor system (Theorem 3.3, ascend). If it is expanding, then set $(\Gamma, \mathcal{G}) = (\Gamma', \mathcal{G}_2)$ and $f = f'$. If it is not expanding, nonsurjectivity of ϕ implies the nonexpanding part corresponds to a proper $[\phi]$-invariant free factor system $\mathcal{G}$ that contains $\mathcal{G}_2$. Collapsing the nonexpanding part will produce a free splitting $(\Gamma, \mathcal{G})$ and expanding relative immersion $f\colon (\Gamma, \mathcal{G}) \to (\Gamma, \mathcal{G})$ whose periodic vertex groups in $\mathcal{G}$ correspond to a larger proper $[\phi]$-fixed free factor system (Proposition 3.2, extend). Either way, it follows that f is the canonical representative for ϕ and we are done.

Remark. The sketch is still not entirely correct as we swept some technicalities under the rug. The representative f_1 need not be an immersion — it could also be *eventually degenerate*! For example, this could happen if the image of ϕ is contained in a proper free factor. Furthermore, as the induction step involves passing to proper free factor systems, the whole proof should be done with *disconnected* free splittings in mind. Thus it is a bit tricky to define the appropriate notion of injectivity and nonsurjectivity in this generality.

Acknowledgments: I would like to thank Alan Logan and the MAXIMALS seminar for pushing me to make my previous existence theorem algorithmic. Logan's comments helped improve the clarity in the paper. I am also grateful to the Max Planck Institute for Mathematics for hosting and supporting me.

Contents

Preparation

A *(connected resp.) topological graph* $(\Gamma, \star)$ is a finite (connected resp.) 1-dimensional CW-complex Γ along with a distinguished vertex (0-cell) for each connected component of Γ; the distinguished vertices $\star \subset \Gamma^{(0)}$ are its *basepoints*; topological graphs are allowed to be *degenerate*, i.e. a finite set of points. A rose is a connected graph with exactly one vertex. All topological graphs will have basepoints even though we will suppress the basepoints from the notation. Each edge (1-cell) consists of two half-edges (or ends) and the *topological tangent space* $T_v\Gamma$ at a vertex $v \in \Gamma^{(0)}$ is finite set of points indexed by the union of v with the half-edges attached to v. We will abuse notation and refer to a point of $T_v\Gamma$ by either v if it is indexed by the vertex v or ϵ if it is indexed by the half-edge ϵ. Nondegenerate edge-paths in Γ are assumed to be oriented and hence have initial and terminal half-edges. For a connected topological graph Γ, the *fundamental group* $\pi_1(\Gamma)$ is the set of degenerate or reduced (i.e. topologically immersed) based loops with a binary operation given by concatenation and tightening/reduction: $(\sigma_1, \sigma_2) \mapsto [\sigma_1\sigma_2]$; and inverses are given by reversals of orientation: $\sigma \mapsto \overline{\sigma}$.

A *cellular map* $f\colon \Gamma \to \Gamma'$ is a continuous function of topological graphs that sends vertices to vertices and edges to possibly degenerate edge-paths; cellular maps induce *topological derivatives* $df_v\colon T_v\Gamma \to T_{f(v)}\Gamma'$. For a cellular map $f\colon \Gamma \to \Gamma'$, let K be the maximum of the combinatorial length of the edge-path $f(e)$ as e varies over all the edges of Γ. Then f is K-Lipschitz, a fact that will be used throughout the paper. Generally, $K(f)$ will denote a convenient Lipschitz constant for f rather than the infimum. *Simplicial maps* are the 1-Lipschitz cellular maps. A cellular map is *based* if it preserves basepoints.

Remark. We are treating Γ topologically to keep the exposition short; all topological notions used in the paper without definition (e.g. immersion, deformation retraction, homotopy equivalence, etc.) have combinatorial counterparts. For completeness, Stallings' paper [Sta83] shows how to work with Γ combinatorially. See also Serre's book [Ser77] and Kapovich–Weidmann–Miasnikov's paper [KWM05] for a purely combinatorial approach to *graph of groups.*

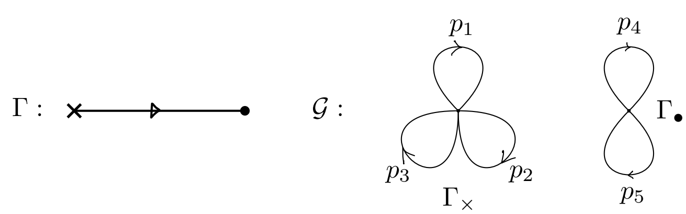

Figure 1: A graph of rank 5; $\times$ marks the basepoint.

Graphs and subgraphs

A graph (of roses) $(\Gamma, \mathcal{G}, g)$ is a triple of two topological graphs and a π_0-bijective map $g\colon \mathcal{G} \to \Gamma^{(0)}$ that labels each vertex $v \in \Gamma^{(0)}$ with a component $\Gamma_v = g^{-1}(v) \subset \mathcal{G}$ that is a rose; the topological graph Γ is the underlying space, the components of $\mathcal{G}$ are the vertex roses, and the map g will now be suppressed from the notation (See Fig. 1). A degenerate graph is a graph whose underlying space is degenerate. A vertex of $(\Gamma, \mathcal{G})$ is a vertex of Γ; a component of $(\Gamma, \mathcal{G})$ is a pair $(G, g^{-1}(G^{(0)}))$ where G is a connected component of Γ. A forest is a graph whose components (trees) have contractible underlying spaces and at most one nondegenerate vertex rose. The tangent space at v is $T_v(\Gamma, \mathcal{G}) = T_v\Gamma \times \pi_1(\Gamma_v)$ and its elements are known as the tangent vectors at v; a tangent vector at v is trivial if its first coordinate is v. A vertex of a graph is bivalent (univalent resp.) if it has exactly two (one resp.) nontrivial tangent vectors; branch points are vertices with at least three nontrivial tangent vectors and natural edges of $(\Gamma, \mathcal{G})$ are maximal edge-paths in Γ whose interior vertices are not branch points nor basepoints in $(\Gamma, \mathcal{G})$. A tight graph is a graph with no univalent vertices except possibly at basepoints.

A subgraph of the graph $(\Gamma, \mathcal{G})$ is a pair $(\Gamma', g^{-1}(\Gamma'))$ where $\Gamma' \subset \Gamma$ is a subcomplex; the subgraph is proper if $\Gamma' \neq \Gamma$. The first examples of subgraphs we have seen are components. The core of a graph $\mathrm{core}(\Gamma, \mathcal{G})$ is a subgraph whose underlying space Γ' is a minimal deformation retract of Γ relative to the vertices with nondegenerate vertex roses; the core is unique unless $(\Gamma, \mathcal{G})$ is a forest with degenerate vertex roses. Closely related to the core, the tightening of a graph $\mathrm{tight}(\Gamma, \mathcal{G})$ is the subgraph whose underlying space Γ' is a minimal deformation retract of Γ relative to the basepoints and vertices with nondegenerate vertex roses. Unlike components and tightenings, core subgraphs need not contain basepoints of the ambient graph and will always be considered without basepoints. For example, a natural edge of a core subgraph can be a concatenation of two natural edges of the original graph if a basepoint is bivalent. All branch points of a tight graph are contained in the graph's core. A branch point of a tight graph is exceptional if it is not a branch point of the core. A natural edge of a tight graph is inner if it is contained in the graph's core. Each component of a tight graph has at most one exceptional branch point.

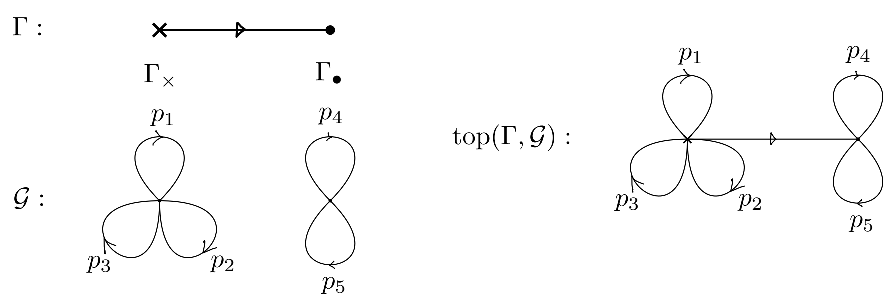

Figure 2: The blow-up of a graph.

Blow-ups and collapses

The next couple of paragraphs introduce some standard operations that can be applied to graphs. The blow-up $\mathrm{top}(\Gamma, \mathcal{G})$ of a graph $(\Gamma, \mathcal{G})$ is the topological graph formed by identifying each vertex $v \in \Gamma$ with the basepoint of the vertex rose Γ_v — the basepoints of $\mathrm{top}(\Gamma, \mathcal{G})$ are the images of the basepoints of Γ. Note that $\mathcal{G}$ is a subcomplex of $\mathrm{top}(\Gamma, \mathcal{G})$ that contains all its vertices (See Fig. 2). Conversely, for any topological graph Γ, define $\mathrm{graph}(\Gamma)$ to be the graph formed by labelling every vertex of Γ with a singleton. Subdivision of a graph is a graph obtained by subdividing the underlying space and extending the vertex roses with singletons that cover the new vertices. This operation can be reversed by *forgetting bivalent vertices*.

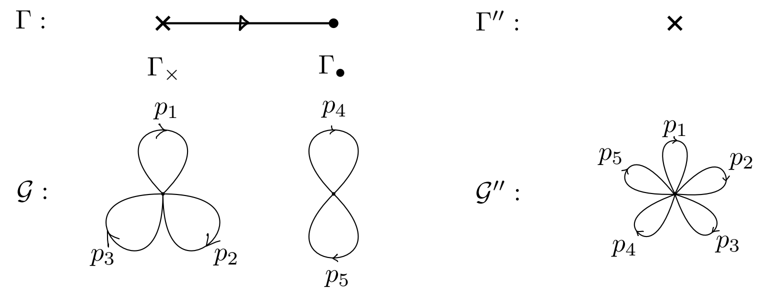

Figure 3: The collapse of a graph with $(\Gamma', \mathcal{G}') = (\Gamma, \mathcal{G})$.

Let $(\Gamma', \mathcal{G}')$ be a subgraph of $(\Gamma, \mathcal{G})$ and $T' \subset \Gamma'$ be a maximal subcomplex with contractible components (topological forest) — each component of Γ' contains exactly one component of T'. Since $\mathcal{G}'$ consists of roses, T' is identified with a maximal topological forest in $\mathrm{top}(\Gamma', \mathcal{G}')$. The collapse of $(\Gamma', \mathcal{G}', T')$ in $(\Gamma, \mathcal{G})$:

1. replace Γ' with the union $\Gamma' \cup \Gamma^{(0)}$ and set $\mathcal{G}' = \mathcal{G}$;

2. let $\Gamma'' = \Gamma/\Gamma'$ be the quotient space of Γ where each component of Γ' is *collapsed* to a vertex, i.e. $\Gamma''^{(0)} = \pi_0(\Gamma')$, $\Gamma''^{(1)} \setminus \Gamma''^{(0)} = \Gamma \setminus \Gamma'$, and the basepoints of Γ'' are the components of Γ' that contain the basepoints of Γ;

3. define $\mathcal{G}'' = \mathrm{top}(\Gamma', \mathcal{G}')/T'$ similarly and for any component v'' of Γ', let $\mathcal{V}''$ be the corresponding component in $\mathrm{top}(\Gamma', \mathcal{G}')/T'$ — the components of $\mathcal{G}''$ are roses;

4. the collapse is the graph $(\Gamma'', \mathcal{G}'')$ with $g'' : \mathcal{G}'' \to \Gamma''^{(0)}$ defined by $g''(\mathcal{V}'') = v''$; by construction, there is a T'-induced based homotopy equivalence $\mathrm{top}(\Gamma, \mathcal{G}) \to \mathrm{top}(\Gamma'', \mathcal{G}'')$.

Up to *graph isomorphism* (defined later), the collapse does not depend on the topological forest T' but, for our purposes, the T'-induced based homotopy equivalence does; nevertheless, we will usually suppress T' and simply say *"collapse* $(\Gamma', \mathcal{G}')$*."* Under certain circumstances (e.g. Γ' is a topological forest), the maximal topological forest $T' \subset \Gamma'$ is unique and the induced based homotopy equivalence is canonical. Finally, $(\Gamma', \mathcal{G}') = (\Gamma, \mathcal{G})$ if and only if the collapsed subgraph is a degenerate graph. Also note that collapsing the subgraph graph$(\mathcal{G})$ in graph(top$(\Gamma, \mathcal{G})$) recovers the graph $(\Gamma, \mathcal{G})$. In Fig. 3, the subgraph is the whole graph and so the collapse is degenerate.

Converse to the collapse construction, suppose $\gamma \colon \mathcal{G} \to \mathrm{top}(\Gamma', \mathcal{G}')$ is a based homotopy equivalence for some graph $(\Gamma', \mathcal{G}')$, then we can consider a labelling of each vertex $v \in \Gamma^{(0)}$ by the component $C'_v \subset \Gamma'$ whose blow-up contains the image $\gamma(\Gamma_v)$; in turn, we can use these labels to define a blow-up $\mathrm{top}_\gamma(\Gamma, \Gamma')$ of the underlying space Γ. The partial blow-up of $(\Gamma, \mathcal{G})$ relative to γ is the graph $\mathrm{rel}_\gamma(\Gamma, \mathcal{G}) = (\mathrm{top}_\gamma(\Gamma, \Gamma'), \mathcal{G}')$ where each vertex $v' \in \mathrm{top}_\gamma(\Gamma, \Gamma')^{(0)}$ is a vertex of Γ' labelled by $\Gamma'_{v'} \subset \mathcal{G}'$. By construction, $(\Gamma', \mathcal{G}')$ is a subgraph of the relative blow-up $\mathrm{rel}_\gamma(\Gamma, \mathcal{G})$ (See Fig. 4), γ can be extended to based homotopy equivalence $\overline{\gamma} \colon \mathrm{top}(\Gamma, \mathcal{G}) \to \mathrm{top}(\mathrm{rel}_\gamma(\Gamma, \mathcal{G}))$, and collapsing $(\Gamma', \mathcal{G}')$ in $\mathrm{rel}_\gamma(\Gamma, \mathcal{G})$ recovers the graph $(\Gamma, \mathcal{G})$ up to a graph isomorphism (defined shortly).

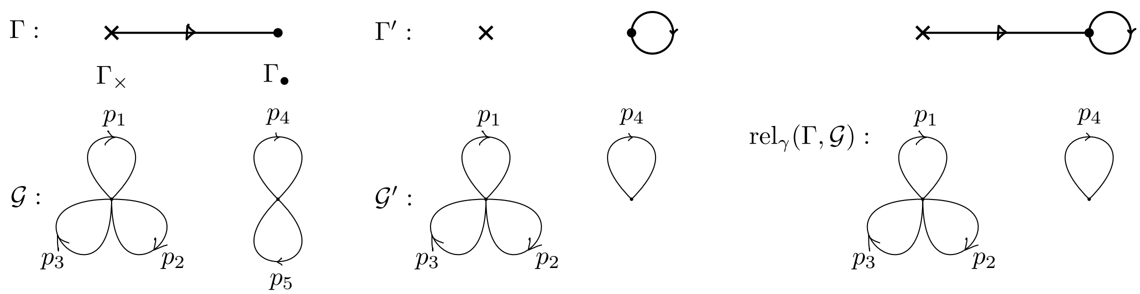

Figure 4: A partial blow-up of a graph.

Graph maps and immersions

A graph map $(f, \Psi) \colon (\Gamma, \mathcal{G}) \to (\Gamma', \mathcal{G}')$ is the following data:

1. cellular maps $f \colon \Gamma \to \Gamma'$ and $\Psi \colon \mathcal{G} \to \mathcal{G}'$ satisfying $g' \circ \Psi = f \circ g$; and
2. for each (oriented) edge e of Γ, a sequence of based loops $\rho(e)_i \in \pi_1(\Gamma'_{v_i})$, where $(v_i)_{i \geq 0}$ is the sequence of vertices along the edge-path $f(e)$.

Unless otherwise stated, graph maps will be based, i.e. the cellular map f is based.

An edge e of Γ is pretrivial if $f(e) \in \Gamma'^{(0)}$; the graph map is degenerate if $f(\Gamma) \subset \Gamma'^{(0)}$, or equivalently, $f(\Gamma)$ is contained in the basepoints of Γ', and it is eventually degenerate if $f^k(\Gamma) \subset \Gamma'^{(0)}$ for some $k \geq 1$. A graph map (f, Ψ) is K-Lipschitz (simplicial resp.) if f is K-Lipschitz (simplicial resp.). We shall occasionally consider unbased graph maps by looking at restrictions of graph maps to subgraphs that have no basepoints.

Graph maps (f, Ψ) induce based cellular maps $\mathrm{top}(f, \Psi)\colon \mathrm{top}(\Gamma, \mathcal{G}) \to \mathrm{top}(\Gamma', \mathcal{G}')$. Conversely, based cellular maps f induce graph maps $\mathrm{graph}(f)\colon \mathrm{graph}(\Gamma) \to \mathrm{graph}(\Gamma')$. A graph map (f, Ψ) is π_1-nonsurjective if $\mathrm{top}(f, \Psi)$ is π_1-nonsurjective, i.e. some based loop in $\mathrm{top}(\Gamma, \mathcal{G})$ is not homotopic rel. basepoints to the $\mathrm{top}(f, \Psi)$-image of a based loop. A homotopy equivalence is a graph map (f, Ψ) whose blow-up $\mathrm{top}(f, \Psi)$ is a based homotopy equivalence. A graph isomorphism is a homotopy equivalence (f, Ψ) whose underlying map f is a simplicial homeomorphism. On the other hand, we will abuse terminology a bit and say (f, Ψ) is π_1-injective if the restriction of $\mathrm{top}(f, \Psi)$ to each component is π_1-injective, i.e. no based loop in $\mathrm{top}(\Gamma, \mathcal{G})$ is homotopic rel. basepoints to the $\mathrm{top}(f, \Psi)$-image of two distinct reduced based loops in the same component of $\mathrm{top}(\Gamma, \mathcal{G})$.

The tightening of a graph has an induced homotopy equivalence whose blow-up is a *deformation retraction*. The collapse of a subgraph $(\Gamma', \mathcal{G}')$ in $(\Gamma, \mathcal{G})$ has a T'-induced homotopy equivalence $(\Gamma, \mathcal{G}) \to (\Gamma'', \mathcal{G}'')$. Conversely, let $\gamma\colon \mathrm{graph}(\mathcal{G}) \to (\Gamma', \mathcal{G}')$ be a homotopy equivalence and $\mathrm{rel}_\gamma(\Gamma, \mathcal{G})$ be the partial blow-up of $(\Gamma, \mathcal{G})$ relative to $\mathrm{top}(\gamma)$. Then $(\Gamma', \mathcal{G}')$ is a subgraph of $\mathrm{rel}_\gamma(\Gamma, \mathcal{G})$ by construction and there is an induced homotopy equivalence $\overline{\gamma}\colon \mathrm{graph}(\mathrm{top}(\Gamma, \mathcal{G})) \to \mathrm{rel}_\gamma(\Gamma, \mathcal{G})$ that extends γ; recall that $\mathrm{graph}(\mathcal{G})$ and $(\Gamma', \mathcal{G}')$ are subgraphs of $\mathrm{graph}(\mathrm{top}(\Gamma, \mathcal{G}))$ and $\mathrm{rel}_\gamma(\Gamma, \mathcal{G})$ respectively.

For any graph map $(f, \Psi)\colon (\Gamma, \mathcal{G}) \to (\Gamma', \mathcal{G}')$, there is an induced derivative at $v \in \Gamma^{(0)}$, $d(f, \Psi)_v\colon T_v(\Gamma, \mathcal{G}) \to T_{f(v)}(\Gamma', \mathcal{G}')$, given by $(\epsilon, \sigma_v) \mapsto (df_v(\epsilon), [\Psi(\sigma_v)\rho(\epsilon)])$, where

- $\rho(\epsilon)$ is trivial when $\epsilon = v$, and
- $\rho(\epsilon) = \rho(e)_0$ when ϵ is the initial half-edge of the oriented edge e.

A graph map (f, Ψ) of tight graphs is natural if $\mathrm{top}(f, \Psi)$ maps natural edges to degenerate or reduced edge-paths and (f, Ψ) maps branch points to branch points with possibly one exception: an exceptional branch point of $(\Gamma, \mathcal{G})$ can be mapped to a bivalent basepoint of $(\Gamma', \mathcal{G}')$. Note that if a natural graph map (f, Ψ) does not map any bivalent basepoint to a univalent basepoint, then it will map cores to cores and the restriction to the cores, denoted $\mathrm{core}(f, \Psi)$, is an unbased graph map.

An immersion is a graph map (f, Ψ) with no pretrivial edges and whose derivative maps $d(f, \Psi)_v$ are injective for all $v \in \Gamma^{(0)}$. Immersions of tight graphs are π_1-injective natural graph maps; π_1-injective graph maps defined on degenerate graphs are vacuously immersions. Although tangent spaces are infinite at vertices with nondegenerate vertex roses, it is a finite check to test injectivity of the derivative maps: if $\epsilon_1 \neq \epsilon_2$ but $df_v(\epsilon_1) = df_v(\epsilon_1)$, then test whether $[\rho(\epsilon_1)\rho(\epsilon_2)^{-1}] \in \Psi_*(\pi_1(\Gamma_v))$, i.e. $\rho(\epsilon_1)\rho(\epsilon_2)^{-1}$ is homotopic rel. basepoints to the Ψ-image of a based loop in Γ_v.

Automorphic expansions

Now consider a graph map $(f, \Psi)\colon (\Gamma, \mathcal{G}) \to (\Gamma, \mathcal{G})$ from a graph to itself. A subgraph of $(\Gamma, \mathcal{G})$ is (f, Ψ)-invariant if the underlying subcomplex is f-invariant. The stable subgraph for (f, Ψ) is the invariant subgraph that consists of f-periodic vertices and edges. An <u>expansion</u> is an immersion whose stable subgraph is degenerate. A graph map (f, Ψ) is <u>automorphic</u> if Ψ restricts to a based homotopy equivalence of the Ψ-periodic components of $\mathcal{G}$. Note that an automorphic graph map restricts to an unbased graph isomorphism on its stable subgraph. The graph map (f, Ψ) <u>permutes basepoints</u> if f is π_0-bijective. The main result of the paper is an algorithm constructing an automorphic expansion *homotopic* to a π_1-injective graph map that permutes basepoints. Fig. 5 is an illustration of an automorphic expansion.

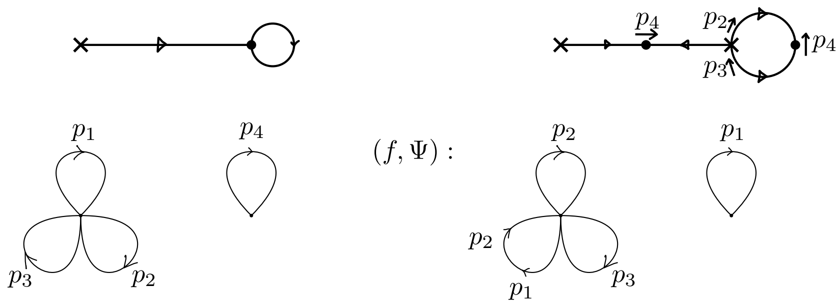

Figure 5: An automorphic expansion on a graph of rank 5. The graph is on the left and the expansion is on the right. Note that the map's image is contained in a proper subgraph.

We will say graph maps (f_1, Ψ_1) and (f_2, Ψ_2) are <u>homotopic</u> via a homotopy equivalence $\alpha\colon (\Gamma_1, \mathcal{G}_1) \to (\Gamma_2, \mathcal{G}_2)$ if $\mathrm{top}((f_2, \Psi_2) \circ \alpha)$ and $\mathrm{top}(\alpha \circ (f_1, \Psi_1))$ are homotopic rel. basepoints. Collapsing an (f, Ψ)-invariant subgraph induces a graph map (f'', Ψ'') that is homotopic to (f, Ψ) via the T'-induced homotopy equivalence. Although we omit the details, we note that the derivative maps of the induced graph map (f'', Ψ'') also depends on the choice of marked point for each component of the collapsed subgraph. Conversely, let $\gamma\colon \mathrm{graph}(\mathcal{G}) \to (\Gamma', \mathcal{G}')$ be a homotopy equivalence and suppose $(f', \Psi')\colon (\Gamma', \mathcal{G}') \to (\Gamma', \mathcal{G}')$ and $\mathrm{graph}(\Psi)$ are homotopic via γ. The partial blow-up of $(\Gamma, \mathcal{G})$ relative to $\mathrm{top}(\gamma)$ induces a graph map $\mathrm{rel}_\gamma(f, \Psi)$ with invariant subgraph $(\Gamma', \mathcal{G}')$ and the induced graph map is natural if (f, Ψ) and (f', Ψ') are natural. The restriction of $\mathrm{rel}_\gamma(f, \Psi)$ to $(\Gamma', \mathcal{G}')$ is (f', Ψ') and $\mathrm{rel}_\gamma(f, \Psi)$ is homotopic to $\mathrm{graph}(\mathrm{top}(f, \Psi))$ via the homotopy equivalence $\overline{\gamma}$ that extends γ. Furthermore, $\mathrm{rel}_\gamma(f, \Psi)$ induces the graph map (f, Ψ) up to a graph isomorphism after collapsing $(\Gamma', \mathcal{G}')$ in $\mathrm{rel}_\gamma(\Gamma, \mathcal{G})$.

Construction

1 Outline

The following algorithm is the main result. We start by outlining the steps in the algorithm using three key steps (Theorems 3.1 and 3.3, Proposition 3.2) as black boxes.

The algorithm for constructing automorphic expansions.

Input: A π_1-injective cellular map $\psi : G \to G$ of roses that permutes basepoints.
Output:

- a graph $(\Delta, \mathcal{D})$, a homotopy equivalence $\alpha\colon \mathrm{graph}(G) \to (\Delta, \mathcal{D})$, and
- an automorphic expansion $p\colon (\Delta, \mathcal{D}) \to (\Delta, \mathcal{D})$ homotopic to $\mathrm{graph}(\psi)$ via α.

  The graph $(\Delta, \mathcal{D})$ is degenerate if and only if ψ is a based homotopy equivalence.

Outline of the algorithm.

Start by checking if ψ is a based homotopy equivalence. If it is, then set

$$(\Delta, \mathcal{D}) = (\pi_0(G), G), \quad \alpha = \pi, \quad p = (\pi_0(\psi), \psi),$$

where $\pi\colon \mathrm{graph}(G) \to (\Delta, \mathcal{D})$ is the evident collapse map and we are done.

Otherwise, ψ is not a based homotopy equivalence. Start with $G_0 = G$ and $\psi_0 = \psi$, then enter the *descent* loop:

1. Given a π_1-injective cellular map $\psi_m\colon G_m \to G_m$ of roses that permutes basepoints but is not a based homotopy equivalence,
2. use the descending algorithm in Theorem 3.1 with ψ_m as input to construct:
   - a tight graph $(\Gamma_{m+1}, \mathcal{G}_{m+1})$ with a nondegenerate core,
   - a homotopy equivalence $\gamma_m\colon \mathrm{graph}(G_m) \to (\Gamma_{m+1}, \mathcal{G}_{m+1})$,
   - a natural graph map $(f_{m+1}, \Psi_{m+1})\colon (\Gamma_{m+1}, \mathcal{G}_{m+1}) \to (\Gamma_{m+1}, \mathcal{G}_{m+1})$ that is homotopic to $\mathrm{graph}(\psi_m)$ via γ_m, and
   - either (f_{m+1}, Ψ_{m+1}) is an immersion or $\mathrm{core}(f_{m+1}, \Psi_{m+1})$ is eventually degenerate.
3. Save the data (f_{m+1}, Ψ_{m+1}) and γ_m for later use.
4. Let G_{m+1} be the Ψ_{m+1}-periodic components of $\mathcal{G}_{m+1}$ and $\psi_{m+1} = \Psi_{m+1}|_{G_{m+1}}$.
5. If ψ_{m+1} is a based homotopy equivalence, exit the loop.
6. Otherwise, restart the loop with ψ_{m+1}.

As each $(\Gamma_m, \mathcal{G}_m)$ has a nondegenerate core, the *complexities* of $\mathcal{G}_m$ are strictly decreasing and the loop will stop after n iterations for some positive $n \leq 2 \cdot \mathrm{rank}(G_0) - 1$. For each positive $m \leq n$, descent produces:

- a tight graph $(\Gamma_m, \mathcal{G}_m)$ with a nondegenerate core,

- a natural graph map $(f_m, \Psi_m)\colon (\Gamma_m, \mathcal{G}_m) \to (\Gamma_m, \mathcal{G}_m)$ that either is an immersion or has an eventually degenerate core, and
- a homotopy equivalence $\gamma_{m-1}\colon \mathrm{graph}(G_{m-1}) \to (\Gamma_m, \mathcal{G}_m)$ such that (f_m, Ψ_m) is homotopic to $\mathrm{graph}(\psi_{m-1})$ via γ_{m-1}, where G_{m-1} is the union of the Ψ_{m-1}-periodic components of $\mathcal{G}_{m-1}$ and $\psi_{m-1} = \Psi_{m-1}|_{G_{m-1}}$ (if $m \geq 2$).

By design, descent stops the first moment it encounters a based homotopy equivalence. Thus (f_n, Ψ_n) is automorphic and each π_1-injective graph map (f_m, Ψ_m) permutes basepoints but is π_1-nonsurjective for $1 \leq m \leq n$. By Lemma 2.3, the graph map (f_n, Ψ_n) is an immersion. In summary, (f_n, Ψ_n) is a π_1-nonsurjective automorphic immersion that permutes basepoints. This concludes the intermediate step of the algorithm.

The graphs, graph maps, and homotopy equivalences produced by the descent will be accessible in the next loop. Start with $m = n$, $(\Omega_n, \mathcal{O}_n) = (\Gamma_n, \mathcal{G}_n)$, $q_n = (f_n, \Psi_n)$, and ω_n the collapse map $\mathrm{graph}(\mathrm{top}(\Gamma_n, \mathcal{G}_n)) \to (\Gamma_n, \mathcal{G}_n)$, then enter the *ascent* loop:

1. Given
   - an index m (for accessing the sequences produced by descent),
   - a nondegenerate graph $(\Omega_m, \mathcal{O}_m)$,
   - a homotopy equivalence $\omega_m\colon \mathrm{graph}(\mathrm{top}(\Gamma_m, \mathcal{G}_m)) \to (\Omega_m, \mathcal{O}_m)$, and
   - an automorphic immersion $q_m\colon (\Omega_m, \mathcal{O}_m) \to (\Omega_m, \mathcal{O}_m)$ that is homotopic to $\mathrm{graph}(\mathrm{top}(f_m, \Psi_m))$ via ω_m — in particular, q_m permutes basepoints and is π_1-nonsurjective,
2. use the extending algorithm in Proposition 3.2 with q_m as input to construct:
   - a nondegenerate graph $(\Delta_m, \mathcal{D}_m)$,
   - a homotopy equivalence $\delta_m\colon (\Omega_m, \mathcal{O}_m) \to (\Delta_m, \mathcal{D}_m)$, and
   - an automorphic expansion $p_m\colon (\Delta_m, \mathcal{D}_m) \to (\Delta_m, \mathcal{D}_m)$ that is homotopic to q_m via δ_m.
3. Set $\alpha_m = \delta_m \circ \omega_m \circ \mathrm{graph}(\mathrm{top}(\gamma_{m-1}))\colon \mathrm{graph}(G_{m-1}) \to (\Delta_m, \mathcal{D}_m)$.

   Note that p_m is homotopic to $\mathrm{graph}(\psi_{m-1})$ via α_m.
4. If $m = 1$, save the data p_m and α_m then exit the loop.
5. Otherwise, recall that either (f_{m-1}, Ψ_{m-1}) is an immersion or $\mathrm{core}(f_{m-1}, \Psi_{m-1})$ is eventually degenerate by the descent construction. Use the ascending algorithm in Theorem 3.3 with p_m, α_m, and (f_{m-1}, Ψ_{m-1}) as inputs to construct:
   - a nondegenerate graph $(\Omega_{m-1}, \mathcal{O}_{m-1})$,

- a homotopy equivalence $\omega_{m-1}\colon \mathrm{graph}(\mathrm{top}(\Gamma_{m-1}, \mathcal{G}_{m-1})) \to (\Omega_{m-1}, \mathcal{O}_{m-1})$,
- and an automorphic immersion $q_{m-1}\colon (\Omega_{m-1}, \mathcal{O}_{m-1}) \to (\Omega_{m-1}, \mathcal{O}_{m-1})$ that is homotopic to $\mathrm{graph}(\mathrm{top}(f_{m-1}, \Psi_{m-1}))$ via ω_{m-1}.

6. Restart the loop with $m-1$, $(\Omega_{m-1}, \mathcal{O}_{m-1})$, ω_{m-1}, and q_{m-1}.

In the end, ascent produces the output for the algorithm:

- a nondegenerate graph $(\Delta_1, \mathcal{D}_1)$,
- a homotopy equivalence $\alpha_1\colon \mathrm{graph}(G_0) \to (\Delta_1, \mathcal{D}_1)$, and
- an automorphic expansion $p_1\colon (\Delta_1, \mathcal{D}_1) \to (\Delta_1, \mathcal{D}_1)$ homotopic to $\mathrm{graph}(\psi_0)$ via α_1.

This concludes the algorithm's outline. □

2 Stallings folds and bounded cancellation

The most important tool in our algorithm is *Stallings factorization* [Sta83]: Stallings showed that any cellular map $f\colon \Gamma \to \Gamma'$ can be algorithmically factored as $f = \iota \circ \gamma$ where ι is a simplicial immersion and γ is a π_1-surjective cellular map — precisely, a composition of pretrivial edge collapses and Stallings *folds*. Furthermore, the maps ι and γ are unique up to simplicial homeomorphism. The cellular map γ is a homotopy equivalence if and only if f is π_1-injective; if f is π_1-injective, then the simplicial immersion ι is a homeomorphism if and only if f is a homotopy equivalence. We now slightly adapt this to work for graph maps.

Lemma 2.1. *Any graph map $(f, \Psi)\colon (\Gamma, \mathcal{G}) \to (\Gamma', \mathcal{G}')$ can be algorithmically factored as $(f, \Psi) = \iota \circ \gamma$ where $\iota\colon S(f, \Psi) \to (\Gamma', \mathcal{G}')$ is a simplicial immersion and $\gamma\colon (\Gamma, \mathcal{G}) \to S(f, \Psi)$ a π_1-surjective graph map. Furthermore, the graph $S(f, \Psi)$ and graph maps ι and γ are unique up to graph isomorphism.*

The graph map γ is a homotopy equivalence if and only if (f, Ψ) is π_1-injective; if (f, Ψ) is π_1-injective, then the simplicial immersion ι is a graph isomorphism if and only if (f, Ψ) is a homotopy equivalence.

The graph $S(f, \Psi)$ will be known as the Stallings graph for (f, Ψ).

Proof. Apply Stallings theorem to factor $\mathrm{top}(f, \Psi) = \iota' \circ \gamma'$ where $\iota'\colon S \to \mathrm{top}(\Gamma', \mathcal{G}')$ is a based simplicial immersion and $\gamma'\colon \mathrm{top}(\Gamma, \mathcal{G}) \to S$ a π_1-surjective based cellular map. Define the subcomplex $\mathcal{S} \subset S$ to be the ι'-preimage of $\mathcal{G}'$. Turning everything into graphs gives $\mathrm{graph}(\mathrm{top}(f, \Psi)) = \mathrm{graph}(\iota') \circ \mathrm{graph}(\gamma')$ where $\mathrm{graph}(\iota')$ is a simplicial immersion and $\mathrm{graph}(\gamma')$ a π_1-surjective graph map.

By construction, $\mathrm{graph}(\mathrm{top}(f, \Psi))$ maps $\mathrm{graph}(\mathcal{G})$ to $\mathrm{graph}(\mathcal{G}')$. Then $\mathrm{graph}(\gamma')$ maps $\mathrm{graph}(\mathcal{G})$ to $\mathrm{graph}(\mathcal{S})$. Collapse $\mathrm{graph}(\mathcal{G})$ in $\mathrm{graph}(\mathrm{top}(\Gamma, \mathcal{G}))$, $\mathrm{graph}(\mathcal{S})$ in $\mathrm{graph}(S)$, and graph($\mathcal{G}'$) in graph(top($\Gamma', \mathcal{G}'$)). This recovers the graphs $(\Gamma, \mathcal{G})$ and $(\Gamma', \mathcal{G}')$ and induces a simplicial immersion $\iota\colon S(f, \Psi) \to (\Gamma', \mathcal{G}')$ (we omit the details here as well) and π_1-surjective graph map $\gamma\colon (\Gamma, \mathcal{G}) \to S(f, \Psi)$, where the graph $S(f, \Psi)$ is the collapse of graph($\mathcal{S}$) in graph(S).

For uniqueness, suppose $(f, \Psi) = \iota'' \circ \gamma''$ where $\iota''\colon S'(f, \Psi) \to (\Gamma', \mathcal{G}')$ is a simplicial immersion and $\gamma''\colon (\Gamma, \mathcal{G}) \to S'(f, \Psi)$ a π_1-surjective graph map. Then $\mathrm{top}(\iota'') = \iota_1 \circ \gamma_1$ where $\iota_1\colon S' \to \mathrm{top}(\Gamma', \mathcal{G}')$ is a based simplicial immersion and $\gamma_1\colon \mathrm{top}(S'(f, \Psi)) \to S'$ a π_1-surjective based cellular map. By uniqueness of Stallings factorization, there is a based simplicial homeomorphism $h\colon S' \to S$ such that $\iota_1 = \iota' \circ h$ and $\gamma' = h \circ \gamma_1 \circ \mathrm{top}(\gamma'')$. Set $\mathcal{S}' \subset S'$ to be the ι_1-preimage of $\mathcal{G}'$. Since ι'' is an immersion, collapsing $\mathcal{S}'$ in S' recovers $S'(f, \Psi)$ up to a graph isomorphism. As $\iota_1 = \iota' \circ h$ and h is a based simplicial homeomorphism, we get that $h(\mathcal{S}') = \mathcal{S}$. So collapsing $\mathcal{S}'$ in S' and $\mathcal{S}$ in S induces a graph isomorphism $\bar{h}\colon S'(f, \Psi) \to S(f, \Psi)$ with $\iota'' = \iota \circ \bar{h}$ and $\gamma = \bar{h} \circ \gamma''$. □

For the most part, we will be interested in the case $(\Gamma', \mathcal{G}') = (\Gamma, \mathcal{G})$ and computing the factorizations of iterates $(f, \Psi)^k$. The next lemma is about based cellular maps, or equivalently, graph maps on graphs with degenerate vertex spaces.

Lemma 2.2. *Let $f\colon \Gamma \to \Gamma$ be a π_1-injective based cellular map that permutes basepoints. If f is not a based homotopy equivalence, then the length of the longest inner natural edge of* $\mathrm{tight}(S(f^k))$ *is unbounded as $k \to \infty$.*

Proof. Suppose $f\colon \Gamma \to \Gamma$ is π_1-injective based cellular map that permutes basepoints and the length of the longest inner natural edge of $\mathrm{tight}(S(f^k))$ was uniformly bounded for all $k \geq 1$. We want to show that f is a based homotopy equivalence. Let $f^k = \iota_k \circ \gamma_k$ be the Stallings factorization for $k \geq 1$. Since f is π_1-injective, γ_k is a based homotopy equivalence, and, as f permutes basepoints, ι_k is π_0-bijective for $k \geq 1$. The number of inner natural edges of $\mathrm{tight}(S(f^k))$ is bounded above by $3 \cdot \mathrm{rank}(\Gamma) - 2$. Thus there is a uniform bound on the number of edges in $\mathrm{core}(S(f^k))$ for $k \geq 1$. So the sequence $\mathrm{core}(\iota_k)$ is eventually periodic, i.e. there are integers $m > n \geq 1$ and a simplicial homeomorphism $h\colon \mathrm{core}(S(f^m)) \to \mathrm{core}(S(f^n))$ such that $\mathrm{core}(\iota_m) = \mathrm{core}(\iota_n) \circ h$.

Find the factorizations $\gamma_n \circ f^{m-n} = \iota' \circ \gamma'$ and $\gamma_1 \circ f^{m-n-1} = \iota'' \circ \gamma''$. On the other hand, $f^m = f^n \circ f^{m-n} = \iota_n \circ \iota' \circ \gamma'$, so uniqueness of factorization implies, up to based simplicial homeomorphism, $\iota_n \circ \iota' = \iota_m$ and $\gamma' = \gamma_m$. So $\mathrm{core}(\iota_n) \circ \mathrm{core}(\iota') = \mathrm{core}(\iota_m)$. Recall that $\mathrm{core}(\iota_m) = \mathrm{core}(\iota_n) \circ h$ where h is a simplicial homeomorphism. First observation: a π_0-bijective simplicial immersion from a graph to itself is a simplicial homeomorphism. Hence, $\mathrm{core}(\iota') \circ h^{-1}$ must be a simplicial homeomorphism. This implies $\mathrm{core}(\iota')$ is a simplicial homeomorphism too. Second observation: a based simplicial immersion whose restriction to the core is a homeomorphism is in fact a based simplicial embedding onto a *deformation retract*. Therefore, ι' is a based simplicial embedding onto a deformation retract.

Since $\gamma' = \gamma_m$ is a based homotopy equivalence, $\gamma_n \circ f^{m-n} = \iota' \circ \gamma'$ is a based homotopy equivalence. As γ_n is a based homotopy equivalence too, so is f^{m-n}. Again by uniqueness of factorization, we may take $\iota_1 \circ \iota'' = \iota_{m-n}$ and $\gamma'' = \gamma_{m-n}$ up to based simplicial homeomorphism. In particular, $\iota_1 \circ \iota''$ is a based simplicial homeomorphism (since f^{m-n} is a based homotopy equivalence) and, consequently, so is ι_1 (and ι''). Therefore, $f = \iota_1 \circ \gamma_1$ is a based homotopy equivalence as γ_1 is a based homotopy equivalence. □

We will now make an observation about graph maps that permute basepoints and restrict to based homotopy equivalences on the periodic vertex roses.

Lemma 2.3. *Suppose $(\Gamma, \mathcal{G})$ has a nondegenerate core and $(f, \Psi) \colon (\Gamma, \mathcal{G}) \to (\Gamma, \mathcal{G})$ is an automorphic graph map that permutes basepoints and preserves the core. If* $\mathrm{core}(f, \Psi)$ *is eventually degenerate, then (f, Ψ) is not π_1-injective.*

Proof. Suppose $\mathrm{core}(f, \Psi)^k$ is degenerate for some $k \geq 1$. Since $\mathrm{core}(f, \Psi)$ permutes components, each component of $\mathrm{core}(\Gamma, \mathcal{G})$ has exactly one periodic vertex and $\mathrm{top}(\mathrm{core}(f, \Psi)^k)$ can be considered a cellular map $\mathrm{top}(\mathrm{core}(\Gamma, \mathcal{G})) \to G$ where G is the union of Ψ-periodic vertex roses labelling vertices of $\mathrm{core}(\Gamma, \mathcal{G})$. Factor $\mathrm{top}(\mathrm{core}(f, \Psi)^k) = \iota \circ \gamma$, then ι is a simplicial homeomorphism since $\mathrm{top}(\mathrm{core}(f, \Psi)^k)$ restricts to a homotopy equivalence of G by the automorphic assumption.

If $\mathrm{top}(\mathrm{core}(f, \Psi)^k)$ were π_1-injective, then ι being a simplicial homeomorphism would imply $\mathrm{top}(\mathrm{core}(f, \Psi)^k)$ were a homotopy equivalence. But as $(\Gamma, \mathcal{G})$ has a nondegenerate core, $\mathrm{rank}(G) < \mathrm{rank}(\mathrm{top}(\Gamma, \mathcal{G}))$ and so $\mathrm{top}(\mathrm{core}(f, \Psi)^k)$ is not π_1-injective. Therefore, $\mathrm{core}(f, \Psi)$ and (f, Ψ) are not π_1-injective. □

The next lemma, also known as the bounded cancellation lemma, will be used extensively in this paper. For an edge-path p in a topological graph Γ, $[p]$ denotes the reduced edge-path that is homotopic to p rel. endpoints.

Lemma 2.4 (Bounded cancellation)**.** *Let $g \colon (\Gamma, \mathcal{G}) \to (\Gamma', \mathcal{G}')$ be a π_1-injective graph map. Then there is a computable constant C such that, for any natural edge-path decomposition $p_1 \cdot p_2$ of a reduced path in* $\mathrm{top}(\Gamma, \mathcal{G})$*, the reduction (rel. endpoints) of the edge-path concatenation* $[\mathrm{top}(g)(p_1)][\mathrm{top}(g)(p_2)]$ *to* $[\mathrm{top}(g)(p_1)\mathrm{top}(g)(p_2)]$ *involves cancelling a subpath with length $\leq C$ in* $[\mathrm{top}(g)(p_1)]$ *and* $[\mathrm{top}(g)(p_2)]$.

The following proof is due to Bestvina-Feighn-Handel [BFH97, Lemma 3.1].

Proof. Factor $\mathrm{top}(g) = \iota \circ \gamma \circ g_0$ into a pretrivial edge collapse and subdivision g_0, a composition of $r \geq 0$ folds $\gamma = g_r \circ \cdots \circ g_1$, and a based simplicial immersion ι. The collapse, subdivision, and immersion have cancellation constants 0 while each fold has cancellation constant 1 by π_1-injectivity. Thus we may choose $C = r$. □

Although the lemma gives a recipe for computing a cancellation constant, $C(g)$ will generally denote an a priori computed constant that may be different from that produced by applying the recipe to g. The main idea is that some operations can increase the size of a graph while keeping the cancellation constant unchanged. If done appropriately, this allows us to promote graph maps to immersions. Our first application of bounded cancellation along these lines is a sufficient condition for a natural graph map to be an immersion.

Lemma 2.5. *Suppose $(f, \Psi)\colon (\Gamma, \mathcal{G}) \to (\Gamma', \mathcal{G}')$ is a π_1-injective natural graph map with cancellation constant $C(f, \Psi)$. If $(\Gamma, \mathcal{G})$ has no pretrivial edges and every natural edge of $(\Gamma', \mathcal{G}')$ is longer than $C(f, \Psi)$, then (f, Ψ) is an immersion.*

Proof. For a contradiction, suppose $d(f, \Psi)_v$ is not injective at the branch point $v \in \Gamma^{(0)}$. By assumption, (f, Ψ) has no pretrivial edges, so $d(f, \Psi)_v(\epsilon_1, \sigma_1) = d(f, \Psi)_v(\epsilon_2, \sigma_2) = (\epsilon, \sigma)$ for some distinct nontrivial tangent vectors $(\epsilon_1, \sigma_1), (\epsilon_2, \sigma_2)$ at v and (ϵ, σ) at $f(v)$; let $\epsilon_1, \epsilon_2, \epsilon$ be respective initial half-edges of oriented natural edges e_1, e_2, e.

Note that $\bar{e}_1 \cdot [\overline{\sigma_1}\sigma_2]e_2$ is a reduced natural edge-path decomposition in $\mathrm{top}(\Gamma, \mathcal{G})$, but

$$[\mathrm{top}(f, \Psi)(\bar{e}_1)][\mathrm{top}(f, \Psi)(\overline{\sigma_1}\sigma_2 e_2)] = u\bar{e}\overline{\rho(\epsilon_1)}\,[\Psi(\overline{\sigma_1}\sigma_2)\rho(\epsilon_2)]\,ew$$

for some reduced edge-paths u, w in $\mathrm{top}(\Gamma', \mathcal{G}')$. The fact that $[\mathrm{top}(f, \Psi)(e_1)]$ has initial segment $\rho(\epsilon_1)e$ follows from (f, Ψ) being natural. Same reasoning goes for the second piece of the concatenation. By assumption, $\sigma = \Psi(\sigma_1)\rho(\epsilon_1) = \Psi(\sigma_2)\rho(\epsilon_2)$, so the reduction of the above edge-path concatenation will involve cancelling a subpath containing e. Yet we assumed the natural edge e is longer than $C(f, \Psi)$, a contradiction. $\square$

The next lemma bounds in terms of a cancellation constant how close π_1-injective graph maps are to being natural.

Lemma 2.6. *Suppose $(f, \Psi)\colon (\Gamma, \mathcal{G}) \to (\Gamma', \mathcal{G}')$ is a π_1-injective graph map of tight graphs with cancellation constant $C = C(f, \Psi)$. Then (f, Ψ) maps branch points to the C-neighborhood of branch points with possibly one exception: some exceptional branch point are mapped to the C-neighborhood of bivalent basepoints.*

Proof. Set $C = C(f, \Psi)$. If $(\Gamma', \mathcal{G}')$ is the C-neighborhood of its branch points, then there is nothing to prove. Suppose ν is a vertex in Γ' whose distance to the nearest branch point is $> C$. We need to show that ν is not the f-image of any branch point of $\mathrm{core}(\Gamma, \mathcal{G})$. Set ϵ_1 to be an oriented half-edges originating from ν and let $\bar{\epsilon}_1$ be the same half-edge with opposite orientation. Since ν is not a branch point, its vertex rose is degenerate. As (f, Ψ) is π_1-injective, all vertices in the preimage $f^{-1}(\nu)$ have degenerate vertex roses.

Suppose v is a branch point of $\mathrm{core}(\Gamma, \mathcal{G})$ and $f(v) = \nu$. As v is a branch point with a degenerate vertex rose, there are at least three distinct oriented half-edges in the core originating from v: e_1, e_2, and e_3. Let p_{12} be an oriented reduced edge-path in $\mathrm{top}(\Gamma, \mathcal{G})$ that starts and ends with e_1 and $\overline{e_2}$ respectively and define p_{23} similarly. Set $p_{13} = [p_{12}p_{23}]$ and $p'_{13} = p_{12}\overline{p_{23}}$. These reduced paths exist because the three oriented half-edges are in the core. See Figure 6 for an illustration. Although the paths are loops, we still treat them as paths, i.e. tightening is done rel. the endpoints. Without loss of generality, assume $[\mathrm{top}(f, \Psi)(p_{12})]$ starts with ϵ_1.

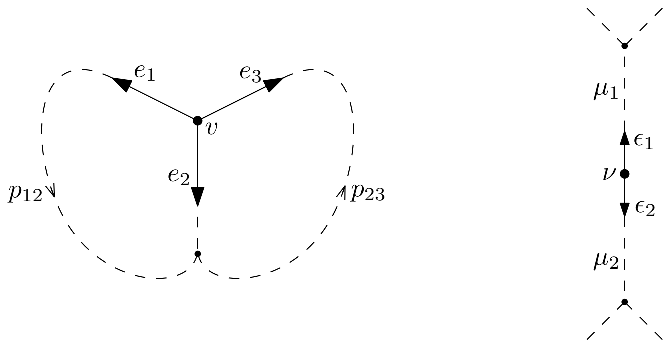

Figure 6: Schematic for paths p_{12}, p_{23}, p_{13}, and p'_{13}. The path p_{13} starts with e_1 follows the dashed path and ends with $\overline{e_3}$. The path p'_{13} is the "figure 8" path traced by p_{12} then $\overline{p_{23}}$. The figure on the right assumes the point ν is in the core but this is not always the case.

If $[\mathrm{top}(f, \Psi)(p_{12})]$ ends with $\bar{\epsilon}_1$, then $[\mathrm{top}(f, \Psi)(p_{12})] = \mu_1 \rho \overline{\mu_1}$, where μ_1 is an extension of ϵ_1 to an embedded path in $\mathrm{top}(\Gamma', \mathcal{G}')$ from ν to a branch point and ρ is a reduced nontrivial loop. By hypothesis, μ_1 is longer than C. Since p_{12} starts and ends with e_1 and $\overline{e_2}$ respectively, the natural edge-path decomposition $p_{12} \cdot p_{12}$ of the immersed path leads to the concatenation of reduced paths $[\mathrm{top}(f, \Psi)(p_{12})][\mathrm{top}(f, \Psi)(p_{12})]$ with $\overline{\mu_1}\mu_1$ as a subpath, violating bounded cancellation. So we may assume ν is bivalent and $[\mathrm{top}(f, \Psi)(p_{12})]$ starts and ends with ϵ_1 and $\bar{\epsilon}_2$, where ϵ_1, ϵ_2 are the distinct oriented half-edges originating from ν.

If $[\mathrm{top}(f, \Psi)(p_{23})]$ starts and ends with ϵ_2 and $\bar{\epsilon}_1$, then $[\mathrm{top}(f, \Psi)(p_{13})]$, the reduction of $[\mathrm{top}(f, \Psi)(p_{12})\mathrm{top}(f, \Psi)(p_{23})]$, starts and ends with ϵ_1 and $\bar{\epsilon}_1$ respectively, which violates bounded cancellation for the same reason given in the previous paragraph. Similarly, if $[\mathrm{top}(f, \Psi)(p_{23})]$ starts and ends with ϵ_1 and $\bar{\epsilon}_2$, we rule out this possibility by considering $[\mathrm{top}(f, \Psi)(p'_{13})]$. We have ruled out all cases, and therefore, branch points of $\mathrm{core}(\Gamma, \mathcal{G})$ are mapped to the C-neighborhood of branch points of $(\Gamma', \mathcal{G}')$.

For the exceptional case, let ν be a vertex in Γ' whose distance to nearest branch point or bivalent basepoint is $> C$. Suppose v is an exceptional branch point of $(\Gamma, \mathcal{G})$ and $f(v) = \nu$. The argument is essentially the same as before. There is a unique oriented half-edge e_* not in the core originating from v and exactly two distinct oriented half-edges e_1, e_2 in the core originating from v. Let p_{12} be defined as before and p_* be the oriented natural edge in $\mathrm{top}(\Gamma, \mathcal{G})$ originating from v and ending at the basepoint $*$. By the argument above, we may assume ν is bivalent and $[\mathrm{top}(f, \Psi)(p_{12})]$ starts and ends with ϵ_1 and $\bar{\epsilon}_2$. Without loss of generality, the reduced path $[\mathrm{top}(f, \Psi)(p_*)]$ starts with ϵ_1. So the natural edge-path decomposition $\overline{p_*} \cdot p_{12}$ leads to a violation of the bounded cancellation lemma. Therefore, an exceptional branch point $(\Gamma, \mathcal{G})$ is mapped to the C-neighborhood of a branch point or bivalent basepoint of $(\Gamma', \mathcal{G}')$. $\square$

Let $q\colon (\Gamma, \mathcal{G}) \to (\Gamma, \mathcal{G})$ be a π_1-injective graph map. For $k \geq 1$, factor $q^k = \iota_k \circ \gamma_k$ into a simplicial immersion $\iota_k\colon S(q^k) \to (\Gamma, \mathcal{G})$ and a homotopy equivalence $\gamma_k\colon (\Gamma, \mathcal{G}) \to S(q^k)$. Set $q_1 = \gamma_1 \circ \iota_1\colon S(q) \to S(q)$ then $\iota_1 \circ q_1 = q \circ \iota_1$ and $q_1 \circ \gamma_1 = \gamma_1 \circ q$. Inductively assume the π_1-injective graph map $q_k\colon S(q^k) \to S(q^k)$ satisfies $\iota_k \circ q_k = q \circ \iota_k$ and $q_k \circ \gamma_k = \gamma_k \circ q$. Factor $q_k = \iota' \circ \gamma'$ into a simplicial immersion $\iota'\colon S' \to S(q^k)$ and a homotopy equivalence $\gamma'\colon S(q^k) \to S'$. We now observe that:

$$\begin{aligned} \iota_k \circ \iota' \circ \gamma' \circ \gamma_k &= \iota_k \circ q_k \circ \gamma_k \\ &= q \circ \iota_k \circ \gamma_k = q^{k+1}. \end{aligned}$$

By uniqueness of factorization, up to graph isomorphism, we may assume $S' = S(q^{k+1})$, $\gamma' \circ \gamma_k = \gamma_{k+1}$, and $\iota_k \circ \iota' = \iota_{k+1}$. Set $q_{k+1} = \gamma' \circ \iota' : S(q^{k+1}) \to S(q^{k+1})$. By construction, $\iota_{k+1} \circ q_{k+1} = \iota_k \circ \iota' \circ \gamma' \circ \iota' = q \circ \iota_{k+1}$ and $q_{k+1} \circ \gamma_{k+1} = \gamma' \circ \iota' \circ \gamma' \circ \gamma_k = \gamma_{k+1} \circ q$.

The first relation $\iota_k \circ q_k = q \circ \iota_k$ and the fact ι_k is a simplicial immersion implies q_k and q have the same Lipschitz and cancellation constants, i.e. $K(q_k) = K(q)$ and $C(q_k) = C(q)$ for all $k \geq 1$. We apply a deformation retraction to the image of q_k to get a homotopic graph map q_k' that preserves tightenings. This does not change the Lipschitz and cancellation constants. Set $\bar{S}_k = \mathrm{tight}(S(q^k))$ and let $\mathrm{tight}(q_k')\colon \bar{S}_k \to \bar{S}_k$ is the restriction of q_k' to the tightenings.

By Lemma 2.6, $\mathrm{tight}(q_k)$ maps branch points to $C(q)$-neighborhoods of branch points except possibly some exceptional branch points are mapped to $C(q)$-neighborhoods of bivalent basepoints. We can apply a *bivalent homotopy* to get a graph map $\bar{q}_k'\colon \bar{S}_k \to \bar{S}_k$ that maps branch points to branch points except possibly some exceptional branch points to bivalent basepoints.

Let $K(q)$ be the Lipschitz constant for q. By the bound on the necessary homotopy, we can use $K(\bar{q}_k') = K(q) + C(q)$ and $C(\bar{q}_k') = 2C(q)$ as the Lipschitz and cancellation constants for $\bar{q}_k'$. We can then apply a *tightening homotopy* to ensure $\mathrm{top}(\bar{q}_k')$ maps natural edges to reduced edge-paths. This homotopy will not worsen the Lipschitz and cancellation constants and the final graph map $\bar{q}_k$ is natural. Finally, let $\bar{\gamma}_k\colon (\Gamma, \mathcal{G}) \to \bar{S}_k$ be the homotopy equivalence induced by the three homotopies. The second relation $q_k \circ \gamma_k = \gamma_k \circ q$ implies $\bar{q}_k$ and q are homotopic via $\bar{\gamma}_k$. To summarize the preceding discussion:

Proposition 2.7. *Suppose $q\colon (\Gamma, \mathcal{G}) \to (\Gamma, \mathcal{G})$ is a π_1-injective graph map. For any $k \geq 1$, there is an algorithm that finds a tight graph $\bar{S}_k$, a homotopy equivalence $\bar{\gamma}_k\colon (\Gamma, \mathcal{G}) \to \bar{S}_k$, and natural graph map $\bar{q}_k\colon \bar{S}_k \to \bar{S}_k$ such that:*

1. *$\bar{S}_k = \mathrm{tight}(S(q^k))$ is the tightening for the Stallings graph for q^k,*
2. *$\bar{q}_k$ is homotopic to q via $\bar{\gamma}_k$, and*
3. *$K(\bar{q}_k) = K(q) + C(q)$ and $C(\bar{q}_k) = 2C(q)$.*

The crucial point is that the Lipschitz and cancellation constants are independent of k.

3 Key steps

We now give the statements and proofs of the key steps in the outline.

Theorem 3.1 (Descend)**.** *Let $\psi\colon G \to G$ be a π_1-injective based cellular map that permutes basepoints and is not a based homotopy equivalence.*

There is an algorithm that takes ψ as input and constructs a tight graph $(\Gamma, \mathcal{G})$ with a nondegenerate core, a homotopy equivalence $\gamma : \mathrm{graph}(G) \to (\Gamma, \mathcal{G})$, and a natural graph map $(f, \Psi)\colon (\Gamma, \mathcal{G}) \to (\Gamma, \mathcal{G})$ such that:

1. *(f, Ψ) is homotopic to* $\mathrm{graph}(\psi)$ *via γ;*
2. *either (f, Ψ) is an immersion or* $\mathrm{core}(f, \Psi)$ *is eventually degenerate.*

Proof. Suppose $\psi\colon G \to G$ permutes basepoints, is tight, but is not a based homotopy equivalence. Set $N = 3 \cdot \mathrm{rank}(G) - 1$, $K = K(\psi) + C(\psi)$, $C = 2C(\psi)$, and $k = 1$. Enter a loop:

1. Given a positive integer $k \geq 1$, construct $S(\psi^k)$ the Stallings graph for $\mathrm{graph}(\psi^k)$.
2. If an inner natural edge of $\mathrm{tight}(S(\psi^k))$ is longer than $C \cdot K^{N-1}$, save k and exit the loop. Otherwise, restart the loop with $k + 1$.

By hypothesis and Lemma 2.2, the lengths of inner natural edges of $\mathrm{tight}(S(\psi^n))$ are unbounded as $n \to \infty$ and so the loop must eventually stop. Use the algorithm in Proposition 2.7 with $\mathrm{graph}(\psi)$ and k as input to construct a homotopy equivalence $\bar{\gamma}_k\colon \mathrm{graph}(G) \to \bar{S}_k$ and a natural graph map $\bar{f}_k\colon \bar{S}_k \to \bar{S}_k$ homotopic to $\mathrm{graph}(\psi)$ via $\bar{\gamma}_k$, where $\bar{S}_k = \mathrm{tight}(S(\psi^k))$; the graph map $\bar{f}_k$ has Lipschitz and cancellation constants K and C respectively.

Form a directed graph $\mathbb{G}_k$ whose vertices are the natural edges of $\bar{S}_k$ and there is a directed edge $E_i \to E_j$ if $\bar{f}_k(E_i)$ (in the underlying space) contains E_j. Since $\bar{S}_k$ is a tight graph with the same rank as G, $\mathbb{G}_k$ has at most N vertices. Let $\mathcal{L}_0$ be the natural edges of $\bar{S}_k$ longer than $C \cdot K^{N-1}$ and $\mathcal{L}$ the union of $\mathcal{L}_0$ and natural edges on a directed path to $\mathcal{L}_0$ in $\mathbb{G}_k$. Since $\bar{f}_k$ is K-Lipschitz and the shortest directed path in $\mathbb{G}_k$ from a natural edge in $\mathcal{L}$ to $\mathcal{L}_0$ has at most N natural edges on it, every natural edge in $\mathcal{L}$ is longer than C. The natural edges in $\mathcal{L}$ will be *long* and the remaining ones will be *short*.

Set $(\Delta', \mathcal{D}')$ be the subgraph of $\bar{S}_k$ consisting of all the short natural edges and vertices. The subgraph is $\bar{f}_k$-invariant by construction of $\mathcal{L}$. Collapsing $(\Delta', \mathcal{D}')$ in $\bar{S}_k$ produces:

1. a tight graph $(\Gamma, \mathcal{G})$ with natural edges longer C and a nondegenerate core,
2. a T'-induced homotopy equivalence $\gamma'_k\colon \bar{S}_k \to (\Gamma, \mathcal{G})$ for some maximal topological forest $T' \subset \Delta'$, and

3. a T'-induced natural graph map $(f, \Psi)\colon (\Gamma, \mathcal{G}) \to (\Gamma, \mathcal{G})$ with cancellation constant C such that (f, Ψ) is homotopic to $\bar{f}_k$ via γ'_k.

Now set $\gamma = \gamma'_k \circ \bar{\gamma}_k$, then (f, Ψ) is homotopic to graph(ψ) via γ. It remains to ensure that either (f, Ψ) is an immersion or it preserves cores and core(f, Ψ) is eventually degenerate.

By the bounded cancellation lemma, no bivalent basepoint is mapped to a univalent basepoint as all natural edges are longer than the cancellation constant C. In particular, the natural graph map (f, Ψ) preserves cores. If no inner natural edge of $(\Gamma, \mathcal{G})$ (as a vertex in $\mathbb{G}_k$) is part of a directed cycle in $\mathcal{L}$, then the restriction core(f, Ψ) is eventually degenerate.

Otherwise, we can iteratively collapse all pretrivial natural edges of $(\Gamma, \mathcal{G})$ and still get a tight graph with natural edges longer than C and a nondegenerate core. Replace the graph $(\Gamma, \mathcal{G})$, homotopy equivalence γ'_k, and (f, Ψ) with the new data produced by the iterative collapses; this ensures (f, Ψ) has no pretrivial edges. As (f, Ψ) is a natural graph map with no pretrivial edges and the natural edges of $(\Gamma, \mathcal{G})$ are longer than the cancellation constant C, the graph map (f, Ψ) is an immersion by Lemma 2.5. □

Proposition 3.2 (Extend)**.** *Let $q\colon (\Omega, \mathcal{O}) \to (\Omega, \mathcal{O})$ be an automorphic immersion that permutes basepoints on a nondegenerate graph $(\Omega, \mathcal{O})$.*

There is an algorithm that takes a π_1-nonsurjective q as input and constructs a nondegenerate graph $(\Delta, \mathcal{D})$, a homotopy equivalence $\delta\colon (\Omega, \mathcal{O}) \to (\Delta, \mathcal{D})$, and an automorphic expansion $p\colon (\Delta, \mathcal{D}) \to (\Delta, \mathcal{D})$ that is homotopic to q via δ.

Proof. Suppose q is a π_1-nonsurjective automorphic immersion on a nondegenerate graph $(\Omega, \mathcal{O})$. Find the stable subgraph $(\Omega', \mathcal{O}')$ for q. If it is degenerate, then q is an expansion and we are done: set $(\Delta, \mathcal{D}) = (\Omega, \mathcal{O})$, δ the identity map, and $p = q$.

Now assume $(\Omega', \mathcal{O}')$ is not degenerate. Since q is automorphic, it restricts to an unbased graph isomorphism of the stable subgraph $(\Omega', \mathcal{O}')$. As q is π_1-nonsurjective, $(\Omega', \mathcal{O}')$ is proper. So collapsing $(\Omega', \mathcal{O}')$ in $(\Omega, \mathcal{O})$ produces:

1. a nondegenerate graph $(\Delta, \mathcal{D})$,

2. a T'-induced homotopy equivalence $\delta\colon (\Omega, \mathcal{O}) \to (\Delta, \mathcal{D})$ for some maximal topological forest T' in Ω', and

3. a T'-induced automorphic graph map $p\colon (\Delta, \mathcal{D}) \to (\Delta, \mathcal{D})$ homotopic to q via δ and with degenerate stable subgraph.

This process may have produced pretrivial edges. The immersion assumption on q implies pretrivial edges are disjoint from periodic vertices. Iteratively collapsing pretrivial edges produces a nondegenerate graph and induces an automorphic expansion. □

Theorem 3.3 (Ascend)**.** *Let $p\colon (\Delta,\mathcal{D}) \to (\Delta,\mathcal{D})$ be an automorphic expansion that permutes basepoints on a nondegenerate graph $(\Delta,\mathcal{D})$ and $\alpha\colon \mathrm{graph}(G) \to (\Delta,\mathcal{D})$ be a homotopy equivalence. Suppose $(f,\Psi)\colon (\Gamma,\mathcal{G}) \to (\Gamma,\mathcal{G})$ is a natural graph map that either is an immersion or has an eventually degenerate core. Furthermore, assume* $\mathrm{graph}(\Psi|_G)$ *is homotopic to p via α; here G is identified with the union of Ψ-periodic components of $\mathcal{G}$.*

There is an algorithm that takes p, α, and (f,Ψ) as input and constructs a nondegenerate graph $(\Omega,\mathcal{O})$, a homotopy equivalence $\omega\colon \mathrm{graph}(\mathrm{top}(\Gamma,\mathcal{G})) \to (\Omega,\mathcal{O})$, and an automorphic immersion q on $(\Omega,\mathcal{O})$ homotopic to $\mathrm{graph}(\mathrm{top}(f,\Psi))$ *via ω.*

Proof. Let $(\Delta,\mathcal{D})$, p, α, $(\Gamma,\mathcal{G})$, and (f,Ψ) be given as in the hypothesis of the theorem. Assume G is the union of Ψ-periodic components of $\mathcal{G}$, the goal is to extend α to a homotopy equivalence from $\mathrm{graph}(\mathcal{G})$. Note that all components are Ψ-preperiodic since $\mathcal{G}$ had finitely many components.

1. If $G = \mathcal{G}$, then exit the loop.
2. Suppose Γ_v is a component of $\mathcal{G} \setminus G$ and $\Gamma_{f(v)}$ a component of G.
3. Set $\Psi_v = \alpha \circ \mathrm{graph}(\Psi|_{\Gamma_v})\colon \mathrm{graph}(\Gamma_v) \to (\Delta,\mathcal{D})$.
4. Factor $\Psi_v = \iota_v \circ \alpha_v$ into a simplicial immersion $\iota_v\colon S(\Psi_v) \to (\Delta,\mathcal{D})$ and homotopy equivalence $\alpha_v\colon \mathrm{graph}(\Gamma_v) \to S(\Psi_v)$. Enlarge $(\Delta,\mathcal{D})$ to include (disjoint union) $S(\Psi_v)$, p to include ι_v, G to include Γ_v, and α to include α_v.
5. The new p is still an automorphic expansion as no periodic edges/vertices were introduced. Furthermore, $\mathrm{graph}(\Psi|_G)$ is still homotopic to p via α. Restart loop.

The loop will stop since all components of $\mathcal{G}$ were preperiodic. Hence now $G = \mathcal{G}$, p is an automorphic expansion on $(\Delta,\mathcal{D})$, and $\mathrm{graph}(\Psi)$ is homotopic to p via α.

Let $(\Omega_0,\mathcal{O}_0) = \mathrm{rel}_\alpha(\Gamma,\mathcal{G})$ be the partial blow-up relative to $\mathrm{top}(\alpha)$, $q_0 = \mathrm{rel}_\alpha(f,\Psi)$ the induced automorphic natural graph map homotopic to $\mathrm{graph}(\mathrm{top}(f,\Psi))$ via the homotopy equivalence $\overline{\alpha}\colon \mathrm{graph}(\mathrm{top}(\Gamma,\mathcal{G})) \to (\Omega_0,\mathcal{O}_0)$ that extends α. Recall that $(\Delta,\mathcal{D})$ is subgraph of $(\Omega_0,\mathcal{O}_0)$ that is q_0-invariant, the restriction of q_0 to $(\Delta,\mathcal{D})$ is p, and q_0 induces (f,Ψ) up to graph isomorphism upon collapsing $(\Delta,\mathcal{D})$. Set $N = 3 \cdot \mathrm{rank}(\mathrm{top}(\Omega_0,\mathcal{O}_0)) - 1$.

Case 1: $\mathbf{core}(\boldsymbol{f},\boldsymbol{\Psi})$ is eventually degenerate. Then $\mathrm{core}(q_0)^k$ has image in the subgraph $(\Delta,\mathcal{D})$ if $k \ge N$. For any $k \ge N$, we can construct the Stallings graph $S(q_0^k)$ for $q_0^k = \iota_k \circ \omega_k$ where ω_k is a homotopy equivalence and ι_k is a simplicial immersion. By the discussion following Lemma 2.6, we can construct $q_k\colon S(q_0^k) \to S(q_0^k)$ such that $\iota_k \circ q_k = q_0 \circ \iota_k$ and $q_k \circ \omega_k = \omega_k \circ q_0$. Since q_0 is automorphic, so is q_k.

As $\mathrm{core}(\iota_k)$ has image in $(\Delta,\mathcal{D})$ and the restriction of q_0 to $(\Delta,\mathcal{D})$ is p, we get $\iota_k \circ q_k|_{\mathrm{core}(S(q_0^k))} = p \circ \mathrm{core}(\iota_k)$. Since ι_k and p are immersions, so is the restriction $q_k|_{\mathrm{core}(S(q_0^k))}$. In particular, the restriction has image in $\mathrm{core}(S(q_0^k))$ and the core is a q_k-invariant subgraph. In fact, the restriction, which we now denote by $\mathrm{core}(q_k)$, is an expansion since ι_k is simplicial and p an expansion. Consequently, the inner natural edges of $\mathrm{tight}(S(q_0^k))$ grow exponentially in length as $k \to \infty$. We now find and fix a $k \geq N$ such that all inner natural edges of $\mathrm{tight}(S(q_0^k))$ are longer than $2C(q_0)$.

We set $(\Omega, \mathcal{O}) = \mathrm{tight}(S(q_0^k))$ and use the deformation retraction $\tau\colon S(q_0^k) \to (\Omega, \mathcal{O})$ to replace q_k with a homotopic rel. $(\Omega, \mathcal{O})$ graph map $q_k'\colon S(q_0^k) \to S(q_0^k)$ that preserves the tightening $(\Omega, \mathcal{O})$. As q_k preserved cores, so does q_k' and $\mathrm{core}(q_k) = \mathrm{core}(q_k')$. Let $q = \mathrm{tight}(q_k')\colon (\Omega, \mathcal{O}) \to (\Omega, \mathcal{O})$ be the restriction of q_k' to $(\Omega, \mathcal{O})$. The graph map q is homotopic to q_k via the deformation retraction τ and homotopic to $\mathrm{graph}(\mathrm{top}(f, \Psi))$ via $\tau \circ \omega_k \circ \overline{\alpha}$. The graph map q may fail to be natural if $(\Omega, \mathcal{O})$ has exceptional branch points. By Lemma 2.6, an exceptional branch point is mapped to the $C(q_0)$-neighborhood of a branch point or bivalent basepoint. After applying an appropriate bivalent homotopy around such branch points, we may assume q is natural and still preserves cores, $C(q) = 2C(q_0)$, and all inner natural edges of $(\Omega, \mathcal{O})$ are longer than $C(q)$. The homotopy produces at most one pretrivial inner natural edge in each component of $(\Omega, \mathcal{O})$. Collapse pretrivial edges if necessary and assume q has no pretrivial edges and $\mathrm{core}(q)$ is once again an automorphic expansion.

It remains to show that q is an immersion. As q is a natural graph map with no pretrivial edges, it is enough to verify injectivity of the derivative at branch points. For a contradiction, suppose dq_v maps two distinct nontrivial tangent vectors at v to the nontrivial tangent vector (ϵ, σ) at $q(v)$; let ϵ be the initial half-edge of oriented natural edge e. Since q is natural and preserves cores, the natural edge e is contained in $\mathrm{core}(\Omega, \mathcal{O})$ and hence longer than $C(q)$. This violates bounded cancellation by the same argument as in the proof of Lemma 2.5.

Case 2: (f, Ψ) is an immersion. Set $K = K(q_0) + C(q_0)$, $C = 2C(q_0)$, and $k = 1$. Enter a loop:

1. Given an integer $k \geq 1$, construct the Stallings graph $S(q_0^k)$ for $q_0^k = \iota_k \circ \omega_k$.

2. As q_0 restricts to the immersion p on the invariant subgraph $(\Delta, \mathcal{D})$ and induces the immersion (f, Ψ) upon collapsing $(\Delta, \mathcal{D})$, the restriction of ω_k to $(\Delta, \mathcal{D})$ is simply a subdivision. Set $(\Delta_k, \mathcal{D}_k)$ to be the ω_k-image of $(\Delta, \mathcal{D})$.

3. The subgraph $(\Delta_k, \mathcal{D}_k)$ has two natural edge structures: 1) its *abstract* natural edges as a graph; 2) natural edges *inherited* from the ambient graph $\mathrm{tight}(S(q_0^k))$. The inherited natural edges partition an abstract natural edge into at most $\frac{2}{3}(N+1)$ segments.

4. If the abstract natural edges in $(\Delta_k, \mathcal{D}_k)$ are all longer than $\frac{2}{3}(N+1) \cdot C \cdot K^{N-1}$, save k and exit the loop. Otherwise, restart the loop with $k+1$.

Since the restriction of q_0 to $(\Delta, \mathcal{D})$ is the expansion p, the lengths of abstract natural edges of $(\Delta_n, \mathcal{D}_n)$ grow exponentially as $n \to \infty$ and the loop must eventually stop. Use the algorithm in Proposition 2.7 with q_0 and k as input to construct a homotopy equivalence $\overline{\omega}_k\colon (\Omega_0, \mathcal{O}_0) \to \bar{S}_k$ and a natural graph map $\overline{q}_k : \bar{S}_k \to \bar{S}_k$ homotopic to q_0 via $\overline{\omega}_k$, where $\bar{S}_k = \mathrm{tight}(S(q_0^k))$; the graph map $\overline{q}_k$ has Lipschitz and cancellation constants K and C respectively. Note that $\overline{q}_k$ is automorphic and $(\Delta_k, \mathcal{D}_k)$ is the $\overline{\omega}_k$-image of $(\Delta, \mathcal{D})$ as well. As p, the restriction of q_0 to $(\Delta, \mathcal{D})$, was already natural, the homotopy between p and the restriction of $\overline{q}_k$ to $(\Delta_k, \mathcal{D}_k)$ via $\overline{\omega}_k$ is supported in the abstract natural edges of $(\Delta_k, \mathcal{D}_k)$.

As in the proof of Theorem 3.1, form a directed graph $\mathbb{G}_k$ whose vertices are the inherited natural edges in $(\Delta_k, \mathcal{D}_k)$ and there is a directed edge $E_i \to E_j$ if $\overline{q}_k(E_i)$ contains E_j; $\mathbb{G}_k$ has at most N vertices. Let $\mathcal{S}_*$ be the inherited natural edges in $(\Delta_k, \mathcal{D}_k)$ shorter than C and $\mathcal{S}$ the union of $\mathcal{S}_*$ and inherited natural edges on a directed path from $\mathcal{S}_*$ in $\mathbb{G}_k$. Since $\overline{q}_k$ is K-Lipschitz and the shortest directed path in $\mathbb{G}_k$ from a natural edge in $\mathcal{S}_*$ to $\mathcal{S}$ has at most N natural edges on it, every natural edge in $\mathcal{S}$ is shorter than $C \cdot K^{N-1}$. The natural edges in $\mathcal{S}$ will be *short*.

By the pigeonhole principle, every abstract natural edge of $(\Delta_k, \mathcal{D}_k)$ has an inherited natural edge longer than $C \cdot K^{N-1}$ and the short inherited natural edges form a $\overline{q}_k$-invariant subforest of $(\Delta_k, \mathcal{D}_k)$. Collapse this subforest in $\bar{S}_k$ to get a graph $(\Omega, \mathcal{O})$ with a subgraph $(\Delta', \mathcal{D}')$ corresponding to $(\Delta_k, \mathcal{D}_k)$. Most importantly, all inherited natural edges in $(\Delta', \mathcal{D}')$ are longer than C. Set $\kappa\colon \bar{S}_k \to (\Omega, \mathcal{O})$ to be the canonical forest collapse homotopy equivalence and $\omega = \kappa \circ \overline{\omega}_k \circ \overline{\alpha}$. The induced automorphic graph map $q\colon (\Omega, \mathcal{O}) \to (\Omega, \mathcal{O})$ has cancellation constant C as well and is homotopic to $\overline{q}_k$ via κ. Then q is homotopic to $\mathrm{graph}(\mathrm{top}(f, \Psi))$ via ω, has a restriction to the q-invariant $(\Delta', \mathcal{D}')$ that is homotopic to p via $\kappa \circ \overline{\omega}_k$, and induces (f, Ψ) up to graph isomorphism upon collapsing $(\Delta', \mathcal{D}')$. Since the short inherited natural edges formed a subforest and the restriction of $\overline{q}_k$ to $(\Delta_k, \mathcal{D}_k)$ was homotopic to the immersion p rel. abstract branch points, the q-pretrivial edges form a subforest of $(\Delta', \mathcal{D}')$. Iteratively collapse pretrivial edges until q has no pretrivial edges.

It remains to show that q is an immersion and, as in the first case, it is enough to verify injectivity of the derivative at branch points. For a contradiction, suppose dq_v maps two distinct nontrivial tangent vectors at v to the nontrivial tangent vector (ϵ, σ) at $q(v)$; let ϵ be initial half-edge of the oriented natural edge e. As q induces the immersion (f, Ψ) upon collapsing $(\Delta', \mathcal{D}')$, e must be contained in $(\Delta', \mathcal{D}')$ and thus longer than C. This violates bounded cancellation and we are done. □

Application

In this section, we list some algorithmic problems that can now be solved. Fix a free group F_k with basis $a_1, \ldots, a_k$. The standard rose is a rose R_k with k petals along with an isomorphism/marking $\mu\colon F_k \to \pi_1(R_k)$ that identifies a_i with the i-th petal. There is a direct way of turning an endomorphism $\phi\colon F_k \to F_k$ into a cellular map $\bar{\phi}\colon R_k \to R_k$ such that $\pi_1(\bar{\phi}) \circ \mu = \mu \circ \phi$.

The algorithm for constructing stable images.
Input: An endomorphism $\phi\colon F_k \to F_k$.
Output: A basis for the stable image $\phi^\infty(F_k) = \bigcap_{i=1}^\infty \phi^i(F_k)$ in terms of $a_1, \ldots, a_k$.

The steps. Let $\phi_i = \phi|_{\phi^i(F)} : \phi^i(F_k) \to \phi^i(F_k)$ for any $i \geq 1$. It follows from the definition that $\phi^\infty(F_k) = \phi_i^\infty(\phi^i(F_k))$. By the Hopfian property, $\psi = \phi_k$ is injective. Let $F_r = \phi^k(F_k)$ with a basis $b_1, \ldots, b_r$; the basis can be algorithmically computed in terms of $a_1, \ldots, a_k$ using Stallings factorization of $\bar\phi^k$. Then $\bar\psi\colon R_r \to R_r$ is a π_1-injective cellular map. The algorithm for finding automorphic expansions takes $\bar\psi$ as input and produces:

- a graph $(\Delta, \mathcal{D})$, a homotopy equivalence $\alpha\colon \mathrm{graph}(R_r) \to (\Delta, \mathcal{D})$, and
- an automorphic expansion $p\colon (\Delta, \mathcal{D}) \to (\Delta, \mathcal{D})$ homotopic to $\mathrm{graph}(\bar\psi)$ via α.

Compute the based homotopy inverse for $\mathrm{top}(\alpha)$ to get a based homotopy equivalence $\beta\colon \mathrm{top}(\Delta, \mathcal{D}) \to R_r$. The restriction of β to the basepoint's vertex rose in $(\Delta, \mathcal{D})$ represents the stable image $\psi^\infty(F_r)$:

- that p is automorphic means the basepoint's vertex rose represents a free factor of F_r contained in $\psi^\infty(F_r)$;
- that p is an expansion means reduced based loops in $\mathrm{top}(\Delta, \mathcal{D})$ with infinitely many iterated based-homotopic preimages are contained in the basepoint's vertex rose.

If the basepoint's vertex rose is degenerate, then the stable image is trivial and we return the trivial element. Otherwise, the homotopy equivalence β and the petals of the vertex rose determine a basis for $\psi^\infty(F_r) = \phi^\infty(F_k)$ in terms of $b_1, \ldots, b_r \in F_r \leq F_k$. $\square$

The algorithm for constructing fixed point subgroups.
Input: An endomorphism $\phi\colon F_k \to F_k$.
Output: A basis for the fixed point subgroup $\mathrm{Fix}(\phi) = \{x \in F_k \ : \ \phi(x) = x\}$.

The steps. Use the previous algorithm to find a basis for the stable image $\phi^\infty(F_k)$. Let $\phi^\infty = \phi|_{\phi^\infty(F_k)}$. By definition, $\mathrm{Fix}(\phi) \leq \phi^\infty(F_k)$ and $\mathrm{Fix}(\phi) = \mathrm{Fix}(\phi^\infty)$. It is evident from the algorithm that $\phi^\infty(F_k)$ is a free factor of $\phi^k(F_k)$ and ϕ^∞ is an automorphism. An algorithm due to Bogopolski–Maslakova [BM16] finds the basis for the fixed point subgroup of free group automorphisms (See also [FH18, Proposition 9.10]). Hence, we can compute a basis for the fixed point subgroup $\mathrm{Fix}(\phi) = \mathrm{Fix}(\phi^\infty)$. $\square$

The algorithm for constructing periodic conjugacy classes.
Input: An injective endomorphism $\phi\colon F_k \to F_k$.
Output: A nontrivial element $x \in F_k$ and integer $i \geq 1$ such that $\phi^i(x)$ is conjugate to x or terminates with no output if no such pair exists.

We will call such conjugacy classes $[x]$ <u>$[\phi]$-periodic</u>.

Sketch of algorithm. The algorithm for finding automorphic expansions produces:

- a graph $(\Delta, \mathcal{D})$, a homotopy equivalence $\alpha\colon \mathrm{graph}(R_k) \to (\Delta, \mathcal{D})$, and
- an automorphic expansion $p\colon (\Delta, \mathcal{D}) \to (\Delta, \mathcal{D})$ homotopic to $\mathrm{graph}(\bar{\phi})$ via α.

As p is an expansion, any nontrivial $[\phi]$-periodic conjugacy class $[x]$ in F_k is represented by a nontrivial reduced (unbased) loop in the periodic vertex roses of $\mathcal{D}$. As p is automorphic, i.e. its restriction to the periodic vertex roses is a homotopy equivalence, the problem of finding a nontrivial $[\phi]$-periodic conjugacy class $[x]$ reduces to the problem of finding a nontrivial homotopically-periodic reduced loop in a homotopy equivalence of roses. An algorithm due to Feighn-Handel [FH18, Corollary 16.4] (dynamically) solves the latter problem (i.e. without invoking *word-hyperbolicity*). $\square$

The algorithm for constructing properly invariant cyclic subgroups.

Input: An injective endomorphism $\phi\colon F_k \to F_k$.
Output: A nontrivial element $x \in F_k$ and integer $i \geq 1$ such that $\phi^i(\langle x \rangle)$ is conjugate to a proper subgroup of $\langle x \rangle$ or terminates with no output if no such pair exists.

We will call such a cyclic group $\langle x \rangle$ properly invariant.

Sketch of algorithm. The algorithm for finding automorphic expansions produces:

- a graph $(\Delta, \mathcal{D})$, a homotopy equivalence $\alpha\colon \mathrm{graph}(R_k) \to (\Delta, \mathcal{D})$, and
- an automorphic expansion $p\colon (\Delta, \mathcal{D}) \to (\Delta, \mathcal{D})$ homotopic to $\mathrm{graph}(\bar{\phi})$ via α.

Since p is an expansion, there is a computable number $L = L(p)$ such that the *iterated pullback* $\hat{\Lambda}_L[\phi]$ (defined below) determines properly invariant cyclic subgroups if nonempty. See [Mut20, Remark 3.12] and also the remark after [Mut21, Proposition 5.5] for details.

Claim. *If ϕ has a properly invariant cyclic subgroup, then $\hat{\Lambda}_i[\phi]$ is nonempty for all $i \geq 1$.*

In particular, if $\hat{\Lambda}_L[\phi]$ is empty, then ϕ has no properly invariant cyclic subgroups and we are done, modulo a proof of the claim. $\square$

Proof of claim. Suppose $x \in F_k$ is nontrivial and $\phi^n(x) = y^{-1}x^d y$ for some $n \geq 1$, $d \geq 2$, and $y \in F_k$. We need to show $\hat{\Lambda}_i[\phi]$ is nonempty for all $i \geq 1$. The iterated pullback $\hat{\Lambda}_i = \hat{\Lambda}_i[\phi]$ is the set of conjugacy classes $[\phi^i(F_k) \cap g\phi^i(F_k)g^{-1}]$ as $[[g]]$ ranges over the $\phi^i(F_k)$-double cosets and $g \notin \phi(F_k)$. Iterated pullbacks can be constructed using topological pullbacks [Sta83]. There is an "inclusion" of $\hat{\Lambda}_{i+1}$ in $\hat{\Lambda}_i$ given by $\phi^{i+1}(F_k) \leq \phi^i(F_k)$; furthermore, the equality $\phi^n(x) = y^{-1}x^d y$ implies $\phi^{2n}(x) = (y\phi^n(y))^{-1}x^{d^2}y\phi^n(y)$. So it is enough to show that $\hat{\Lambda}_j$ is nonempty for some $j \geq n$. Specifically, we will show $\hat{\Lambda}_{2n-m}$ is nonempty for some nonnegative $m < n$.

First define $\psi : F_k \to F_k$ by $u \mapsto y\phi^n(u)y^{-1}$. Observe that $\psi(x) = x^d$ and $x \notin \psi^\infty(F_k)$ since the restriction of ψ to $\psi^\infty(F_k)$ is an automorphism — an automorphism cannot map an element to a proper power. Let $x \in \psi^m(F_k)$ but $x \notin \psi^{m+1}(F_k)$ with $0 \le m$. If $x = \psi^m(z)$ with $z \notin \psi(F_k)$, then the equality $\psi(x) = x^d$ implies $\psi(z) = z^d$ by injectivity of ψ^m. Replace x with z if necessary and assume $x \notin \psi(F_k)$, or equivalently $y^{-1}xy \notin \phi^n(F_k)$. Once again let $y^{-1}xy \in \phi^m(F_k)$ but $y^{-1}xy \notin \phi^{m+1}(F_k)$ with $0 \le m < n$.

If $y^{-1}xy = \phi^m(z)$ with $z \notin \phi(F_k)$, then the equality

$$\phi^{2n}(x) = \phi^n(y)^{-1}y^{-1}xy\phi^n(y)\phi^{2n}(x)\phi^n(y)^{-1}y^{-1}x^{-1}y\phi^n(y)$$

implies (by injectivity of ϕ^m and inequality $m < n$)

$$\phi^{2n-m}(x) = \phi^{n-m}(y)^{-1}z\phi^{n-m}(y)\phi^{2n-m}(x)\phi^{n-m}(y)^{-1}z^{-1}\phi^{n-m}(y);$$

moreover, $g = \phi^{n-m}(y)^{-1}z\phi^{n-m}(y) \notin \phi(F_k)$ as $z \notin \phi(F_k)$ and $m < n$. In particular, $[\phi^{2n-m}(x)]$ is a nontrivial conjugacy class *supported* in $[\phi^{2n-m}(F_k) \cap g\phi^{2n-m}(F_k)g^{-1}]$; the latter is an element of $\hat{\Lambda}_{2n-m}$ since $g \notin \phi(F_k)$ and we are done. $\square$

The algorithm for (dynamically) detecting hyperbolicity.

Input: An injective homomorphism $\phi\colon A \to F$ where $A \le F$ is a free factor.
Output: A correct yes/no answer to whether the HNN extension

$$F*_A = \langle F, t \mid t^{-1}xt = \phi(x)\ \forall x \in A\rangle \text{ is } \textit{word-hyperbolic}.$$

Sketch of algorithm. The proof of [Mut21, Proposition 7.1] is essentially an algorithm for constructing the *canonical invariant free factor system* $\mathcal{F}$ for $[\phi]$. By [Mut21, Theorem 7.5], the HNN extension $F*_A$ is word-hyperbolic if and only if the restriction of $[\phi]$ to $\mathcal{F}$ has neither periodic conjugacy classes nor properly invariant cyclic subgroups. Thus, we can detect word-hyperbolicity by combining the last two (dynamical) algorithms. $\square$

References

[BFH97] Mladen Bestvina, Mark Feighn, and Michael Handel. Laminations, trees, and irreducible automorphisms of free groups. *Geom. Funct. Anal.*, 7(2):215–244, 1997.

[BH92] Mladen Bestvina and Michael Handel. Train tracks and automorphisms of free groups. *Ann. of Math. (2)*, 135(1):1–51, 1992.

[BM16] Oleg Bogopolski and Olga Maslakova. An algorithm for finding a basis of the fixed point subgroup of an automorphism of a free group. *Internat. J. Algebra Comput.*, 26(1):29–67, 2016.

[CL22] Laura Ciobanu and Alan D. Logan. Fixed points and stable images of endomorphisms for the free group of rank two. *J. Algebra*, 591:538–576, 2022.

[FH18] Mark Feighn and Michael Handel. Algorithmic constructions of relative train track maps and CTs. *Groups Geom. Dyn.*, 12(3):1159–1238, 2018.

[IT89] Wilfried Imrich and Edward C. Turner. Endomorphisms of free groups and their fixed points. *Math. Proc. Cambridge Philos. Soc.*, 105(3):421–422, 1989.

[KWM05] Ilya Kapovich, Richard Weidmann, and Alexei Miasnikov. Foldings, graphs of groups and the membership problem. *Internat. J. Algebra Comput.*, 15(1):95–128, 2005.

[Mut20] Jean Pierre Mutanguha. Hyperbolic immersions of free groups. *Groups Geom. Dyn.*, 14(4):1253–1275, 2020.

[Mut21] Jean Pierre Mutanguha. The dynamics and geometry of free group endomorphisms. *Adv. Math.*, 384:107714, 2021.

[Ser77] Jean-Pierre Serre. *Arbres, amalgames,* SL_2. Société Mathématique de France, Paris, 1977. Avec un sommaire anglais, Rédigé avec la collaboration de Hyman Bass, Astérisque, No. 46.

[Sta83] John R. Stallings. Topology of finite graphs. *Invent. Math.*, 71(3):551–565, 1983.

[Tur96] Edward C. Turner. Test words for automorphisms of free groups. *Bull. London Math. Soc.*, 28(3):255–263, 1996.

[Ven10] Enric Ventura. Computing fixed closures in free groups. *Illinois J. Math.*, 54(1):175–186, 2010.