CONJUGACY IN A FAMILY OF FREE-BY-CYCLIC GROUPS

M. R. BRIDSON, T. R. RILEY AND A. W. SALE

Abstract. We analyse the geometry and complexity of the conjugacy problem in a family of free-by-cyclic groups $H_m = F_m \rtimes \mathbb{Z}$ where the defining free-group automorphism is positive and polynomially growing. We prove that the conjugator length function of H_m is linear, and describe polynomial-time solutions to the conjugacy problem and conjugacy search problem in H_m.

2020 Mathematics Subject Classification: 20F65, 20F10

Key words and phrases: free-by-cyclic groups, conjugacy problem, conjugator length

1. introduction

Suppose G is a finitely generated group. The *conjugacy problem* asks for an algorithm that, given any words u and v on the generators and their inverses, decides whether or not these words represent conjugate elements in G. We write $u \sim v$ to denote conjugacy. The *conjugacy search problem* asks for an algorithm that, given a pair of words u and v such that $u \sim v$, will output a word w with $uw = wv$ in G. The *conjugator length function* $\mathrm{CL} : \mathbb{N} \to \mathbb{N}$ quantifies these problems: $\mathrm{CL}(n)$ is the least integer N such that for all words u and v that represent conjugate elements in G and have length $|u| + |v| \leq n$, there is a word w of length at most N such that $uw = wv$ in G. The conjugator length functions of G with respect to different finite generating sets are $\simeq$-equivalent, where $\simeq$ is the equivalence relation that identifies functions $\mathbb{N} \to \mathbb{N}$ that dominate each other modulo affine distortions of their domain and their range. Extensive background on conjugator length can be found in [BRSb].

Fix an integer $m \geq 1$ and let $F = F(a_1, \ldots, a_m)$ be a rank-m free group. Define $\varphi \in \mathrm{Aut}(F)$ by $\varphi(a_i) = a_i a_{i-1}$ for $2 \leq i \leq m$ and $\varphi(a_1) = a_1$. This paper concerns the free-by-cyclic groups

$$H_m \ = \ F \rtimes_\varphi \mathbb{Z} \ = \ \langle a_1, \ldots, a_m, s \mid s^{-1} a_i s = \varphi(a_i) \rangle.$$

The inclusions $H_{m-1} \hookrightarrow H_m$ (excluding a_m) and retractions $H_m \to H_{m-1}$ (killing a_1) will facilitate induction arguments.

The groups H_m have many useful properties and have appeared regularly in the literature. They appear as 'hydra groups' in [BR20, DR13, DER18, Pue16]. Each is the fundamental group of a compact non-positively curved 2-complex built from squares [Sam06]; in particular it is biautomatic and CAT(0). Each can be expressed as a 2-generator 1-relator group,

---

Date: 22 May 2025.

We gratefully acknowledge the financial support of the Royal Society (MRB), the Simons Foundation (TRR–Simons Collaboration Grant 318301), and the National Science Foundation (TRR–GCR2428489 and OIA-2428489).

or as a free-by-cyclic group $F_r \rtimes \mathbb{Z}$ with r arbitrarily large [But07]. H_2 is famous as a 3-manifold group that is not subgroup separable [BKS87]. But, most obviously, these groups H_m serve as natural prototypes for the mapping tori of free-group automorphisms that have maximal polynomial growth [BBMS97, Bri02, CM11, Ger94, Mac02, Mac13, Sam06], and for the most part this is how we shall regard them.

Our main results here are:

Theorem 1. *For all $m \geq 2$, the conjugator length function of H_m satisfies* $\mathrm{CL}(n) \simeq n$.

Theorem 2. *For all $m \geq 2$, there exist algorithms solving the conjugacy problem and the conjugacy search problem of H_m in time polynomial in the sum of the lengths of the input words.*

Our proofs of these theorems are intertwined and are constructive. We first describe in detail an algorithmic procedure that solves the conjugacy problem and the conjugacy search problem of H_m; this is summarized in Section 10. A naïve analysis shows that this algorithm will output a conjugator whose length is bounded by a quadratic function of the lengths of the input words, but a more careful analysis shows that with minor modifications this quadratic bound can be reduced to a linear one—see Remark 10.1.

We regard these results as a significant step towards bounding the complexity of the conjugacy problem and conjugacy search problem in arbitrary free-by-cyclic groups (where the free group has finite rank, which will be a standing assumption throughout our discussion). Free-by-cyclic groups provide a rich and challenging arena for the study of geometric invariants of groups associated with various weak forms of non-positive curvature (as discussed in [BG10], for example). For any free-by-cyclic group, there is an algorithm solving its word problem in polynomial time [Sch08]. There are also algorithms solving the conjugacy problem [BMMV06, BG10], but these do not provide reasonable bounds on time complexity. In particular, it is unknown whether the conjugacy problem and conjugacy search problem can be solved in polynomial time. The results in this paper add weight to the conviction that this is likely.

When a free-by-cyclic group is hyperbolic or its conjugacy problem and conjugacy search problem can be solved in linear time. These are basic examples of a much more general result: there are polynomial-time solutions for all groups which are hyperbolic relative to a finite family of peripheral subgroups in which one can solve the corresponding problems in polynomial time—see [Bum15, EH06, JOR10, O'C]. Free-by-cyclic groups are hyperbolic relative to a finite family of free-by-cyclic subgroups, each of which has the property that the defining automorphism is polynomial-growing—see [BFW19] for history and references. Thus the search for a polynomial time solution to the conjugacy problem reduces to the case where the defining automorphism is polynomial-growing, and Theorem 2 solves this problem for a natural family of prototypes.

It seems reasonable to expect that the conjugator length function of an arbitrary free-by-cyclic group is linear. This is true in the hyperbolic case [BH99, Lys89], but beyond that little is known. However, by appealing to the relative hyperbolicity result mentioned above, one can again reduce to the case where the defining automorphism is polynomially growing, because Sale proved [AS16] that if G is non-degenerately hyperbolic relative to parabolic subgroups P_ω ($\omega \in \Omega$), then $\mathrm{CL}_G(n) \simeq \max\{\mathrm{CL}_{P_\omega}(n) : \omega \in \Omega\} + n$. As in the case of complexity, Theorem 1 assures us that the desired bound $\mathrm{CL}_G(n) \simeq n$ is valid in a natural class of prototypes.

The role that the groups H_m play as prototypes among free-by-cyclic groups is analogous to the role that the *model filiform groups*

$$\Gamma_m = \mathbb{Z}^m \rtimes \mathbb{Z} \ = \ \langle\, a_1, \ldots, a_m, s \mid a_i a_j = a_j a_i \ \forall i, j,\ s^{-1} a_i s = a_i a_{i-1} \forall i \geq 2,\ s^{-1} a_1 s = a_1 \,\rangle,$$

play among (free-abelian)-by-cyclic groups. In [BR] we prove that, in contrast to Theorem 1, the conjugator length function of Γ_m is polynomial of degree m.

The proofs is this paper are largely combinatorial and typically require a delicate analysis of cases. We make heavy use of the notion of 'decomposing reduced words into pieces.' This tool is from [DR13], and can be viewed as a special case of the train-track machinery of [BH92, BFH00, BFH05]. We have favoured using pieces here because they lend themselves well to the detailed study of cancellation in the free group that we need, and to the precise understanding of how words in the free group grow under iteration of the automorphism. Nevertheless, we have structured our proofs with an eye to how they might be adapted to cover more general polynomially growing automorphisms. In particular, we have not relied on any of the alternative ways of viewing $H_m = F \rtimes \mathbb{Z}$ that were discussed earlier. Instead, we consistently view H_m as a semidirect product and work with elements in the form wt^n, where $w \in F$ and $n \in \mathbb{Z}$. From this viewpoint, the complexity of the conjugacy problem in H_m translates into a collection of *twisted conjugacy problems* in F. A benefit of this direct approach is that the outlines of various arguments carry over to the general case.

In a sequel to this paper [BRSa], we will present a different approach to the conjugacy problem in H_m that does rely on one of these alternative perspectives, namely the fact that H_m can be obtained from $\mathbb{Z}^2$ by a sequence of HNN extensions with cyclic amalgamated groups. The more geometric arguments in [BRSa] are framed with an eye to further generalisations.

In the next section we will translate the conjugacy problem in H_m into a suite of twisted conjugacy problems in F and lay out the framework for the rest of this article. It is the analysis of these twisted problems that forms the bulk of what follows. Throughout, we shall write H in place of H_m when there is no danger of ambiguity.

2. Reduction to twisted problems in F

Conjugacy in the free group $F = F(a_1, \ldots, a_m)$ can be fully understood thanks to the following well-known result (e.g. [LS07]).

Lemma 2.1. *If words u and v on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$ represent conjugate elements of F, then there is a word w on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$ which is a concatenation of a prefix of u^{-1} with a suffix of v such that $uw = wv$ in F. (If v is cyclically reduced—that is, vv is reduced—then w need only be a prefix of u^{-1}.)*

Assume u does not represent the identity. Take k to be the maximal integer such that there exists u_0 with $u_0^k = u$ in F. (So u_0 generates the centralizer of u in F.) Then for any such w and u_0,

$$\left\{ W \in F \;\middle|\; uW = Wv \text{ in } F \right\} \ = \ \left\{ u_0^l \;\middle|\; l \in \mathbb{Z} \right\} w.$$

The conjugacy problem for H. *Given words u and v on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}, s^{\pm 1}$ does there exist a word w on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}, s^{\pm 1}$ such that $uw = wv$ in H?*

We will use a standard free-by-cyclic **normal form**: each element in $H = F \rtimes \langle s \rangle$ can be expressed uniquely as $\tilde{u}s^p$ for some reduced word $\tilde{u}$ on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$ and $p \in \mathbb{Z}$.

Suppose we have conjugate elements u and v of H expressed in normal form as $\tilde{u}s^p$ and $\tilde{v}s^q$, respectively. The conjugacy relation $uw = wv$ in H, where w has normal form $\tilde{w}s^r$, implies $p = q$ and amounts to the 'φ-twisted conjugacy relation'

(1) $$\tilde{u}\varphi^{-p}(\tilde{w}) \;=\; \tilde{w}\varphi^{-r}(\tilde{v}) \qquad \text{in the free group } F.$$

This problem is much harder than the conjugacy problem for free groups, although as we will see, in some instances its solution ultimately reduces to Lemma 2.1.

In the instance where $p = q = 0$, the conjugacy problem in H therefore amounts to:

The 0-twisted-conjugacy problem. *Given words $\tilde{u}$, $\tilde{v}$ on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$, do there exist $r \in \mathbb{Z}$ and $\tilde{w} \in F$ such that $\tilde{u}\tilde{w} = \tilde{w}\varphi^{-r}(\tilde{v})$ in F?*

This problem is addressed in Section 5. Proposition 5.1 gives both complexity and conjugator length bounds.

The conjugacy problem in H with $p < 0$ is equivalent to that with $p > 0$ since we can exchange u and v with their inverses. So in place of $p \neq 0$, let us just consider $p = q > 0$.

From $uw = wv$ in H, we get that $u(u^j w) = (u^j w)v$ for all $j \in \mathbb{Z}$, and so there exists a w such that $uw = wv$ in H *and* such that the normal form of w is $\tilde{w}s^r$ for some reduced word $\tilde{w}$ on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$ and some integer r satisfying $0 \leq r < p$.

The Cayley graph of F is a tree, so the geodesics joining 1, $\tilde{u}$, $\tilde{w}$, and $\tilde{u}\varphi^{-p}(\tilde{w})$ form either the '$\mathcal{H}$-configuration' (left) or the 'I-configuration' (right) shown in Figure 1. Accordingly, we can find prefixes u_0, v_0 and suffixes u_1, v_1 of $\tilde{u}, \varphi^{-r}(\tilde{v})$, respectively, and two words x, y, at least one the empty word, such that $\tilde{w} = u_0 x v_0^{-1}$ and $\varphi^{-p}(\tilde{w}) = u_1^{-1} x v_1$, where $\tilde{u} = u_0 y u_1$ and $\varphi^{-r}(\tilde{v}) = v_0 y v_1$ as freely reduced words.

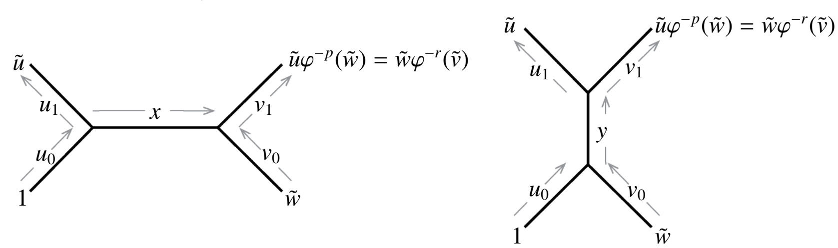

FIGURE 1. The two possibilities for the relative locations of 1, $\tilde{u}$, $\tilde{w}$, and $\tilde{u}\varphi^{-p}(\tilde{w})$ in the Cayley graph of F: the '$\mathcal{H}$-configuration' on the left and the 'I-configuration' on the right.

In the $\mathcal{H}$-configuration, the conjugacy problem amounts to:

The $\mathcal{H}$-twisted conjugacy problem. *Given reduced words $\tilde{u}$, $\tilde{v}$ on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$ and $p > 0$, do there exist $0 \leq r < p$ and words $x, u_0, v_0, u_1, v_1 \in F$ such that $\tilde{u} = u_0 u_1$ and $\varphi^{-r}(\tilde{v}) = v_0 v_1$, as words, and*

$$\varphi^{-p}(u_0 x v_0^{-1}) \;=\; u_1^{-1} x v_1 \;\; in \;\; F?$$

Most of the difficulties and technicalities lie in this problem. Section 6 addresses a special case of the $\mathcal{H}$-twisted conjugacy problem when $\tilde{u}$ is the empty word. As explained there, this amounts to understanding the structure of common prefixes of a word $\tilde{w}$ with its image $\varphi^r(\tilde{w})$, a crucial ingredient in the general case of the $\mathcal{H}$-twisted conjugacy problem. In Section 7, specifically Proposition 7.1, we describe how to use a solution (that is, the integer r and the words x, u_0, u_1, v_0, v_1) to obtain a 'nicer' solution in which x is replaced by a word X whose structure can be described in terms of short chunks that come as subwords of u_0, u_1, v_0, v_1, their inverses and the iterates under powers of φ. This 'chunky' structure enables us, in Section 8, to find a short conjugator and describe a polynomial-time solution to this problem.

Unfortunately, the short conjugator that one obtains from the 'chunky' structure of Proposition 7.1 actually only has a quadratic upper bound in terms of its length in H, relative to those of u and v. Lemma 8.4 describes how to replace one of the chunks with a suitable power of s to obtain a linearly bounded conjugator.

In the I-configuration, x is the empty word and $\tilde{w} = u_0 v_0^{-1}$, and so the conjugacy problem amounts to:

The I-twisted conjugacy problem. *Given reduced words $\tilde{u}$, $\tilde{v}$ on $a_1^{\pm 1}, \dots, a_m^{\pm 1}$ and $p > 0$, do there exist $0 \leq r < p$ and prefixes u_0 of $\tilde{u}$ and v_0 of $\varphi^{-r}(\tilde{v})$ such that*

$$\tilde{u}\varphi^{-p}(u_0 v_0^{-1}) \;=\; u_0 v_0^{-1} \varphi^{-r}(\tilde{v}) \;\; in \;\; F?$$

This problem is easy to solve by an exhaustive search. Indeed, given $\tilde{u}, \tilde{v}$ and p as in the I-twisted conjugacy problem, define $\mathcal{I}$ to be the set of all pairs $(\tilde{w}, r)$, where $0 \leq r < p$, and $\tilde{w}$ is a word of the form UV where U is a prefix of $\tilde{u}$ and V^{-1} is a prefix of $\varphi^{-r}(\tilde{v})$. A solution, if it exists, can be found by applying the solution to the word problem in F to check the validity of each equation $\tilde{u}\varphi^{-p}(\tilde{w}) = \tilde{w}\varphi^{-r}(\tilde{v})$ for each $(\tilde{w}, r) \in \mathcal{I}$.

Like the conjugacy problem, the 0-twisted-conjugacy problem, the $\mathcal{H}$-twisted conjugacy problem, and the I-twisted conjugacy problem all have 'search' variants in which one is given that a collection of integers and words solving the problem exists and is required to exhibit one.

For $g \in H$, let $|g|_H$ denote the length of a shortest word on $\{a_1, \dots, a_m, s\}$ that represents g. If $g \in F$, let $|g|_F$ be the length of a shortest word on $\{a_1, \dots, a_m\}$ that represents g. For a word (not necessarily reduced) w, $\ell(w)$ denotes the number of letters in w.

The following summarises results from Proposition 5.1 and Corollary 8.3, along with the discussion above for the I-twisted case.

Proposition 2.2. *With the notation established above, in H_m, the 0-twisted conjugacy problem, the $\mathcal{H}$-twisted conjugacy problem, and the I-twisted conjugacy problem can each be solved by deterministic algorithms with input $(\tilde{u}, \tilde{v}, p)$ whose running time is bounded by a polynomial in $p + \ell(\tilde{u}) + \ell(\tilde{v})$. And the same is true for the 'search' variants of these problems.*

Given that $|p| \leq \ell(u)$, $\ell(\tilde{u}) \leq C\ell(u)^m$, and $\ell(\tilde{v}) \leq C\ell(v)^m$ for a suitable constant $C > 0$, Theorem 2 follows from Proposition 2.2 as per the above discussion.

Turning to conjugator length, since we want to bound conjugator length in H, bounds pertaining to the three subordinate problems in F will need to be given in terms of the word metric on H rather than on F. The issue of comparing these two metrics is delicate and is the subject of Section 4.

In summary, the article is structured as follows. Section 3 introduces *piece decompositions*, a useful technical tool. Section 4 addresses the distortion of F in H. Section 5 deals with the 0-twisted conjugacy problem. After technical results in Section 6 on common prefixes of words $\tilde{w} \in F$ and their iterates under powers of φ, we handle the $\mathcal{H}$-twisted-conjugacy problem instances in Sections 7 and 8. The conjugator length argument in H is completed in Section 9. Section 10 summarizes how to assemble our results into an algorithm for Theorem 2. In Section 11 we focus on the structure of H_m as an iterated HNN extension and outline an alternative solution to the conjugacy problem.

3. Preliminaries: piece decompositions, definitions, and conventions

3.1. **Some notations, and conventions.**

For a word w on a set of letters, we let $\ell(w)$ denote its length. As mentioned in Section 2, $|\cdot|_H$ and $|\cdot|_F$ denote the word length of an element in H or F respectively, with respect to generating sets $\{a_1, \cdots, a_m, s\}$ and $\{a_1, \cdots, a_m\}$ respectively.

The *rank*, written $\operatorname{rank}(w)$, of a word w on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$ is the maximal i such that $a_i^{\pm 1}$ appears in w. The empty word has rank 0. The rank of $g \in F$ is the rank of the reduced word w representing g.

For a word w, when we write $\varphi(w)$ we mean the reduced word representing $\varphi(w)$ in F.

3.2. **Positivity of φ and φ^{-1}.**

The inverse of φ is

$$\varphi^{-1}(a_i) = \begin{cases} a_{2k}a_{2(k-1)} \cdots a_2 a_1^{-1} a_3^{-1} \cdots a_{2k-1}^{-1} & \text{when } i = 2k, \\[2ex] a_{2k+1}a_{2k-1} \cdots a_1 a_2^{-1} a_4^{-1} \cdots a_{2k}^{-1} & \text{when } i = 2k+1. \end{cases} \tag{2}$$

A useful feature of φ is that it is a *positive automorphism*: whenever $g \in F$ is represented by a positive word, so is $\varphi(g)$. This is not true of φ^{-1}. However φ^{-1} is positive with respect to the basis $b_1, \ldots, b_m$, defined by $b_i = a_i^{(-1)^{i+1}}$ for all i, since

$$\varphi^{-1}(b_i) = \begin{cases} b_{2k-1} \cdots b_3 b_1 b_2 \cdots b_{2(k-1)} b_{2k} & \text{when } i = 2k, \\[2ex] b_{2k+1} b_{2k-1} \cdots b_1 b_2 b_4 \cdots b_{2k} & \text{when } i = 2k+1. \end{cases} \tag{3}$$

3.3. **Pieces and their types.**

We will find it useful to split w into *pieces* that behave well when one takes iterated images under φ. A *rank-i piece* in w is a maximal subword of one of the following four *types*:

$$a_i u, \quad u a_i^{-1}, \quad a_i u a_i^{-1}, \quad u,$$

where u is a (possibly empty) word of rank at most $i-1$. Pieces of the first three types are said to be of *strict rank-i*. Each rank-i word can be expressed as a concatenation of a minimal number of rank-i pieces in a unique manner. We call this *the rank-i decomposition of w* and refer to the pieces involved as the *pieces of w*. We denote the number of these pieces by $|w|_\pi$. For example, $w = a_3 a_2 a_1^{-1} a_3 a_3 a_1 a_3^{-1} a_2 = (a_3 a_2 a_1^{-1})(a_3)(a_3 a_1 a_3^{-1})(a_2)$ is a rank-3 word with $|w|_\pi = 4$.

For $g \in F$ we write $|g|_\pi := |w|_\pi$, where w is a reduced word representing g.

We need the following facts about pieces:

Lemma 3.1. *If a reduced word π is a piece of rank i, then both $\varphi(\pi)$ and $\varphi^{-1}(\pi)$ are also pieces of rank-i and have the same type as π.*

Lemma 3.2. *Let w be a reduced word of rank i. Let $w = \pi_1 \cdots \pi_p$ be its rank-i decomposition. Then there is no cancellation between pieces on applying φ or φ^{-1}—that is, for $k = 1, \ldots, p-1$, the words $\varphi^{\pm 1}(\pi_{k+1})$ and $\varphi^{\pm 1}(\pi_k)$ start and end (respectively) with letters that are not mutual inverses. As a consequence, $\varphi^r(\pi_1) \cdots \varphi^r(\pi_p)$ is freely reduced and is the rank-i decomposition of $\varphi^r(w)$ for all $r \in \mathbb{Z}$.*

We leave the proofs of Lemmas 3.1 and 3.2 as exercises. Very similar observations are made in [DR13].

4. Growth rates and distortion

In order to analyze the φ-twisted conjugacy problem in F we will examine in Section 4.1 how free group elements grow in length on repeated application of φ. Then in Section 4.2 we establish some useful inequalities relating the normal form of $g \in H$ to $|g|_H$.

4.1. Growth rates.

Our next few results lead into Proposition 4.4, which gives a precise estimate of how words grow on repeated applications of $\varphi^{\pm 1}$. (Cf. [Lev09] in which bounds are given, but with the constants depending on the group element.)

Lemma 4.1. $\mathrm{Fix}(\varphi) = \langle a_1,\, a_2 a_1 a_2^{-1} \rangle$.

Proof. That $\mathrm{Fix}(\varphi) \supseteq \langle a_1,\, a_2 a_1 a_2^{-1} \rangle$ is straight-forward. For the reverse inclusion, first observe that by Lemmas 3.1 and 3.2, w is fixed by φ if and only if its pieces are all fixed by φ. So we can focus on the case where w is a single non-empty piece $\pi = a_i^{\delta} u a_i^{-\delta'}$ such that $\pi = \varphi(\pi)$ in F, and $\delta, \delta' \in \{0, 1\}$ are not both zero, and u a reduced word of rank at most $i-1$ with $i \geq 2$. Applying φ to π adds $\delta - \delta'$ to the exponent sum of the a_{i-1} present. But since $\pi = \varphi(\pi)$, the exponent sum of the a_{i-1} in π and $\varphi(\pi)$ must agree, and therefore $\delta = \delta' = 1$ and $\pi = a_i u a_i^{-1}$ (and, in particular, u is non-empty).

It remains to show that $i = 2$. Assume, for contradiction, that $i > 2$. Well, $\pi = \varphi(\pi)$ tells us that $a_i u a_i^{-1} = \varphi(a_i u a_i^{-1}) = a_i a_{i-1} \varphi(u) a_{i-1}^{-1} a_i^{-1}$, and so $\varphi(u) = a_{i-1}^{-1} u a_{i-1}$. Reapplying φ multiple times gives

$$\varphi^r(u) \;=\; \varphi^{r-1}(a_{i-1}^{-1}) \cdots a_{i-1}^{-1} u a_{i-1} \cdots \varphi^{r-1}(a_{i-1}), \quad \text{for } r \geq 1.$$

By Lemma 3.2, the number of pieces p in the rank-$(i-1)$ decompositions of u and $\varphi^r(u)$ are the same for all $r \geq 1$. Hence there is cancellation in our expression for $\varphi^r(u)$. This cancellation can occur only at either end of u. For r large enough, say $r > p+1$, we will need u to completely cancel out under free reduction. In particular, there are some α and β such that

$$1 \;=\; \varphi^{\alpha}(a_{i-1}^{-1}) \ldots a_{i-1}^{-1} u a_{i-1} \ldots \varphi^{\beta}(a_{i-1})$$

and so

$$u \;=\; a_{i-1} \cdots \varphi^{\alpha}(a_{i-1}) \varphi^{\beta}(a_{i-1}^{-1}) \cdots a_{i-1}^{-1}.$$

If $\alpha = \beta$, then we get complete cancellation and $u = 1$, a contradiction. So we assume $\alpha \neq \beta$. We can count the number of pieces p of u by observing the locations of letters $a_{i-1}^{\pm 1}$

(we use here that $i > 2$). We get $p = |u|_\pi = \alpha + \beta + 1$ (the terms $\varphi^\alpha(a_{i-1})\varphi^\beta(a_{i-1}^{-1})$ merge into one piece). We also, from above, have

$$\varphi^r(u) = \varphi^{r-1}(a_{i-1}^{-1}) \cdots \varphi^{\alpha+1}(a_{i-1}^{-1})\varphi^{\beta+1}(a_{i-1}) \cdots \varphi^{r-1}(a_{i-1}),$$

which has to cancel down to p pieces. Since $i > 2$, and $\alpha \neq \beta$, this cancels down to $2r - \alpha - \beta - 3 = 2r - p - 2$ pieces. Since $r > p + 1$, we get a contradiction. □

The following argument is well known (cf. Lemmas 3.5 and 5.6 in [BR09] and Example 3.3 [Bri02]).

Lemma 4.2. *Let $i \geq 2$. There exist $C_i, D_i > 0$ such that for all $r \in \mathbb{Z} \setminus \{0\}$,*

$$C_i\, |r|^{i-1} \;\leq\; |\varphi^r(a_i)|_F \;\leq\; D_i\, |r|^{i-1}.$$

Proof. First we address the case $r > 0$. Since $\varphi^r(a_i)$ is a positive word in this case, its length is the same as its length in the abelianisation of F. With respect to the basis $\{a_1, \ldots, a_m\}$, the action of φ on the abelianisation is via the matrix Φ with ones on the diagonal and immediately above, and zeros elsewhere. Direct calculation yields upper triangular matrices:

$$\Phi = \begin{pmatrix} 1 & 1 & & & & \\ & 1 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & 1 & 1 \\ & & & & & 1 \end{pmatrix}, \qquad \Phi^r = \begin{pmatrix} 1 & \binom{r}{1} & \binom{r}{2} & \cdots & \binom{r}{m-1} \\ & 1 & \binom{r}{1} & \cdots & \binom{r}{m-2} \\ & & \ddots & & \vdots \\ & & & 1 & \binom{r}{1} \\ & & & & 1 \end{pmatrix},$$

where $\binom{r}{j}$ is understood to be 0 for $j > r$.

So

$$|\varphi^r(a_i)|_F = \sum_{j=0}^{i-1} \binom{r}{j},$$

which, as a function of $r \in \mathbb{N}$, is Lipschitz equivalent to $\binom{r}{i-1} \sim r^{i-1}$.

To deal with negative powers, we use the fact that φ^{-1} is a positive automorphism with respect to the basis $\{b_1, \ldots, b_m\}$ described in Section 3.2. With respect to this basis, the action of φ^{-1} on the abelianisation of F is given by

$$\Psi = \begin{pmatrix} 1 & 1 & 1 & \cdots & \cdots & 1 \\ & 1 & 1 & \cdots & \cdots & 1 \\ & & \ddots & & & \vdots \\ & & & \ddots & & \vdots \\ & & & & 1 & 1 \\ & & & & & 1 \end{pmatrix}, \qquad \Psi^r = \begin{pmatrix} 1 & \binom{r}{1} & \binom{r+1}{2} & \cdots & \binom{r+m-2}{m-1} \\ & 1 & \binom{r}{1} & \cdots & \binom{r+m-3}{m-2} \\ & & \ddots & & \vdots \\ & & & 1 & \binom{r}{1} \\ & & & & 1 \end{pmatrix}.$$

So

$$|\varphi^r(a_i)|_F = \sum_{j=0}^{i-1} \binom{r+j-1}{j}$$

which, as a function of $r \in \mathbb{N}$, is Lipschitz equivalent to $\binom{r+i-2}{i-1} \sim r^{i-1}$. □

A straight-forward induction (cf. [DR13, Lemma 7.1]) gives:

Lemma 4.3. *For $i = 2, \ldots, m$,*

$$\varphi^r(a_i) \;=\; \begin{cases} a_i\, a_{i-1}\varphi(a_{i-1})\cdots\varphi^{r-1}(a_{i-1}) & \text{when } r > 0 \\ a_i\varphi^{-1}(a_{i-1})^{-1}\varphi^{-2}(a_{i-1})^{-1}\cdots\varphi^{r}(a_{i-1})^{-1} & \text{when } r < 0. \end{cases}$$

Furthermore, these expressions are reduced words.

One sees that the second expression is reduced by appealing again to the fact φ^{-1} is positive with respect to the basis $\{a_1, a_2^{-1}, a_3, \ldots, a_m^{\pm 1}\}$.

Proposition 4.4. *Suppose $i \in \{2, \ldots, m\}$ and π is a piece of strict rank-i that is not fixed by φ. Then $|\varphi^r(\pi)|_F \sim |r|^{i-1}$. More precisely, for the constants $C_i, D_i > 0$ of Lemma 4.2, for all $r \neq 0$,*

$$|\varphi^r(\pi)|_F \;\leq\; D_i\,|r|^{i-1}\,|\pi|_F$$

and when $|r|^{i-1} \geq \frac{1}{C_i}\,|\pi|_F$,

$$|\varphi^r(\pi)|_F \;\geq\; \left(C_i^{\frac{1}{i-1}}\,|r| - |\pi|_F{}^{\frac{1}{i-1}}\right)^{i-1}.$$

Proof. The upper bound follows via a simple induction argument on the rank i, taking D_i from Lemma 4.2. We therefore focus on the lower bound.

We have $\pi = a_i^{\delta}ua_i^{-\delta'}$ a rank-i piece with $\delta, \delta' \in \{0, 1\}$ not both zero and u a reduced word of rank at most $i-1$. We may assume $\delta = 1$, as otherwise we could replace π by π^{-1}.

Case 1. $\delta' = 0$ and $\pi = a_i u$.

Let $u = \rho_1\cdots\rho_p$ be the piece decomposition of u. Our argument will be that any cancellation between the images of a_i and u under iterated applications of φ or φ^{-1} will occur within the first p applications. After this, there is no further cancellation, so the length of π will eventually have growth rate at least that of a_i under iterated applications of φ or φ^{-1}, which will lead to the required lower bound.

Case 1a. $r > 0$.

If the first letter of u is a_{i-1}, then there is no cancellation between $\varphi^r(a_i)$ and $\varphi^r(u)$ for any $r > 0$, since $\varphi^r(u)$ has first letter a_{i-1} and $\varphi^r(a_i)$ is a positive word.

So suppose, on the other hand, that the first letter of u is not a_{i-1}, and that there is cancellation between $\varphi(a_i)$ and $\varphi(u)$. Then we may write $\rho_1 = va_i^{-\varepsilon}$, for $\varepsilon \in \{0, 1\}$ and v a word of rank at most $i-2$ (if $i = 2$ then v is the empty word). Then $\varphi(a_i v a_{i-1}^{-\varepsilon}) = a_i a_{i-1}\varphi(v)a_{i-2}^{-\varepsilon}a_{i-1}^{-\varepsilon}$ (if $i = 2$, we read a_0 as the empty word). But $\varphi(v)a_{i-2}^{-\varepsilon}$ has rank strictly less than $i-1$, so in order for there to be cancellation between $\varphi(a_i)$ and $\varphi(u)$ we must have $\varepsilon = 1$ and $a_{i-1}\varphi(v)a_{i-2}^{-1}a_{i-1}^{-1} = 1$. That is,

$$\varphi(\pi) \;=\; a_i\varphi(\rho_2)\cdots\varphi(\rho_p).$$

We repeat the argument, and conclude that after $k \leq p$ steps we will reach a situation where we have

$$\varphi^k(\pi) \;=\; a_i\varphi^k(\rho_{k+1})\cdots\varphi^k(\rho_p)$$

and there will be no cancellation between images of a_i and $\varphi^k(\rho_{k+1})$ on further applications of φ. (If $k = p$ we understand that $\rho_{k+1} \cdots \rho_p$ is the empty word.) Hence, for $r \geq k$, we may write

$$\varphi^r(\pi) \;=\; \varphi^{r-k}(a_i)\varphi^r(\rho_{k+1} \cdots \rho_p)$$

where there is no cancellation in the right-hand side. In particular, this gives

$$|\varphi^r(\pi)|_F \;\geq\; \left|\varphi^{r-k}(a_i)\right|_F \;\geq\; C_i(r-k)^{i-1}$$

by Lemma 4.2.

We complete this case by bounding k. Indeed, since $\varphi^k(a_i\rho_1 \cdots \rho_k) = a_i$, we get $\left|\varphi^{-k}(a_i)\right|_F \leq |\pi|_F$. Then an application of Lemma 4.2 gives $C_i k^{i-1} \leq |\pi|_F$, which implies the required lower bound of $|\varphi^r(\pi)|_F$.

Case 1b. Assume $r < 0$.

The argument is broadly similar to Case 1a. The key observation is that there will be no cancellation between images of a_i and u under applications of φ^{-1} whenever the first piece of u is of the form $va_{i-1}^{-\delta}$, for $\delta \in \{0, 1\}$. By induction, the last letter of $\varphi^r(a_i)$ as a reduced word on $\{a_1, \ldots, a_m\}$ is a_{i-1}^{-1} for all $r < 0$. (The base case, $r = -1$, is from (2).) So the only way we can ever have cancellation between $\varphi^r(a_i)$ and $\varphi^r(u)$ is if $\varphi^r(\rho_1)$ begins with a_{i-1}. But since the type of a piece is preserved under applications of φ, if $\rho_1 = va_{i-1}^{-\delta}$, then this will never occur.

So we may assume $\rho_1 = a_{i-1}va_{i-1}^{-\delta}$, for $\delta \in \{0, 1\}$ and v is a word of rank at most $i-2$. Then

$$\varphi^{-1}(a_i\rho_1) \;=\; \varphi^{-1}(a_ia_{i-1})\varphi^{-1}(va_{i-1}^{-\delta}) \;=\; a_i\varphi^{-1}(va_{i-1}^{-\delta}).$$

In particular, if $\varphi^{-1}(va_{i-1}^{-\delta}) \neq 1$ then further applications of φ^{-1} will lead to no cancellation between the images of a_i and of $\varphi^{-1}(va_{i-1}^{-\delta})$, and we can stop. Otherwise $\varphi^{-1}(va_{i-1}^{-\delta}) = 1$, which implies that $va_{i-1}^{-\delta} = 1$, and $\rho_1 = a_{i-1}$. This gives

$$\varphi^{-1}(\pi) \;=\; a_i\varphi^{-1}(\rho_2) \cdots \varphi^{-1}(\rho_p).$$

Repeating this, we find for some $k \leq p$ that

$$\varphi^{-k}(\pi) \;=\; a_i\varphi^{-k}(\rho_{k+1}) \cdots \varphi^{-k}(\rho_p)$$

and there will be no cancellation between images of a_i and $\varphi^{-k}(\rho_{k+1})$ after further applications of φ^{-1}. (If $k = p$ we understand that $\rho_{k+1} \cdots \rho_p$ is the empty word.) As above we then yield, for $r < -k$

$$|\varphi^r(\pi)|_F \;\geq\; \left|\varphi^{r+k}(a_i)\right|_F \;\geq\; C_i(|r|-k)^{i-1}$$

by Lemma 4.2, and $C_ik^{i-1} \leq |\pi|_F$, completing this case.

Case 2. $\delta' = 1$ and $\pi = a_iua_i^{-1}$.

We can apply the arguments from Case 1 to both ends of u and it is not hard to see that the same conclusion is reached. The key point is that if u is completely cancelled out under iterated applications of $\varphi^{\pm 1}$, then the cancellation cannot reach the middle of u simultaneously, meaning that either after several applications of φ we obtain $a_i\varphi^k(a_i^{-1})$, or $\varphi^k(a_i)a_i^{-1}$. The length of these grow as required on further applications of $\varphi^{\pm 1}$. □

4.2. **Tools for handling the distortion.** Recall that for $h \in H$, $|h|_H$ denotes the length of the shortest word on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}, s^{\pm 1}$ representing h in H.

Proposition 4.5. *Suppose $h \in H$ has normal form $\tilde{u}s^p$. If u' is a subword of $\tilde{u}$, then $|u'|_H \le (2m+1)\,|h|_H$.*

Remark 4.6. This proposition does not rely on any special properties of φ: with a change of constant, it holds for any free-by-cyclic group $M = F \rtimes_\psi \mathbb{Z}$. We shall sketch a geometric proof of this more general fact that will allow the reader familiar with van Kampen diagrams to skip the algebraic proof that follows. This geometric argument assumes that the reader is familiar with the use of s-corridors (as used in [BG10], for example).

Let w be a shortest word in the generators $\{a_1, \ldots, a_n, s\}$ that equals h in M, and consider a least-area van Kampen diagram Δ with boundary label $w^{-1}\tilde{u}s^p$. This diagram is a union of its s-corridors and each point along the side of a corridor is a distance at most C from a point along the other side, where C is a constant that depends on ψ.

Let x and y be the endpoints of the arc in $\partial\Delta$ labelled u'. It suffices to argue that x can be connected to y by a path in the 1-skeleton of Δ that has length at most $2pC + |w|$. To construct such a path, observe that every vertex z on the arc of $\partial\Delta$ labelled $\tilde{u}$ can be connected to a vertex on the arc A of $\partial\Delta$ labelled w by crossing at most p of the s-corridors, i.e. the corridors emanating from the arc of $\partial\Delta$ labelled s^p. Thus there is a path α_z of length at most pC from z that ends on A. The desired path from x to y is obtained by following α_x then proceeding along A to the endpoint of α_y before returning along α_y.

To aid the intuition of readers who wish to persist with the algebraic proof, we present an example.

Example 4.7. Suppose h is represented by the word $u = sa_6a_5^{-1}s^{-2}a_5s^2a_3$. Advance the s at the lefthand end through u until it cancels with the first s^{-1}, applying φ^{-1} to the letters $a_i^{\pm 1}$ it passes, to get

$$u_1 \;:=\; (a_6a_4a_2a_1^{-1}a_3^{-1}a_5^{-1})(a_4a_2a_1^{-1}a_3^{-1}a_5^{-1})s^{-1}a_5s^2a_3$$

satisfying $u = u_1$ in H. Then advance the s^{-1} likewise until it cancels with the s to get

$$u_2 \;:=\; (a_6a_4a_2a_1^{-1}a_3^{-1}a_5^{-1})(a_4a_2a_1^{-1}a_3^{-1}a_5^{-1})(a_5a_4)sa_3$$

satisfying $u_1 = u_2$ in H. Then advance the remaining s to the right end to get

$$u_3 \;:=\; (a_6a_4a_2a_1^{-1}a_3^{-1}a_5^{-1})(a_4a_2a_1^{-1}a_3^{-1}a_5^{-1})(a_5a_4)(a_3a_1a_2^{-1})$$

satisfying $u_2 = u_3s$ in H. Then $u = u_3s$ in H, and $\tilde{u}$ is the reduced version of u_3. Suppose $u' = a_3^{-1}a_4a_3$, a subword close to the right-hand end of $\tilde{u}$. Let $u_3' = a_3^{-1}a_5^{-1}a_5a_4a_3$, the subword of u_3 that freely reduces to u'. If we take $u_2' = a_3^{-1}a_5^{-1}a_5a_4$, then $u_3' = u_2'sa_3$; if we take $u_1' = a_3^{-1}a_5^{-1}s^{-1}a_5$, then $u_2' = u_1's$; and if we take $u_0' = s^{-1}a_5$, then $u_1' = a_3^{-1}a_5^{-1}u_0'$. In particular, u_0' is a subword of u, and it is obtained from u_3' by pre- and post-multiplying by a number (bounded by the exponent sum of s in u, and therefore by $|h|_H$ if u is of minimal length) of short words (the images of letters $a_i^{\pm 1}$ under $\varphi^{\pm 1}$, with possibly an s added at the beginning or end).

So there is a word that equals u' in H and whose length can be bounded from above by the length of a subword of u plus the sum of the lengths of these short words. The strategy of the following proof is to bound $|u'|_H$ accordingly.

Proof of Proposition 4.5. Suppose u is a word on $a_1^{\pm1}, \ldots, a_m^{\pm1}, s^{\pm1}$ representing h. On account of the free-by-cyclic structure of H and the fact that $u = \tilde{u}s^p$ in H, there is a sequence of words $u_0, \ldots, u_r$ and integers $p_0, \ldots, p_r$ with the following properties (for all i):

- $r \leq \ell(u)$,
- $u_0 = u$ (as words) and $p_0 = 0$,
- u_r freely reduces to $\tilde{u}$ and $p_r = p$,
- $u = u_i s^{p_i}$ in H (and so the exponent sum of the $s^{\pm1}$ in $u_i s^{p_i}$ is equal to that of u),
- there are letters $x_1, \ldots, x_n \in \left\{a_1, a_1^{-1}, \ldots, a_m, a_m^{-1}\right\}$ (which depend on i) such that, as words, either
  (1) $u_i = \alpha_i s^{\pm1} x_1 \cdots x_n s^{\mp1} \beta_i$ and $u_{i+1} = \alpha_i \varphi^{\mp1}(x_1) \cdots \varphi^{\mp1}(x_n) \beta_i$ and $p_{i+1} = p_i$, or
  (2) $u_i = \alpha_i s^{\pm1} x_1 \cdots x_n$ and $u_{i+1} = \alpha_i \varphi^{\mp1}(x_1) \cdots \varphi^{\mp1}(x_n)$ and $p_{i+1} = p_i \pm 1$,
  for some words α_i and β_i. (In the first case $u_i = u_{i+1}$ in H. In the second, $u_i s^{\mp1} = u_{i+1}$ in H.) The words u_i need not be reduced.

Suppose u'_{i+1} is a subword of u_{i+1}. We claim that there is a subword u'_i of u_i and there are words μ_{i+1} and λ_{i+1} with $\ell(\mu_{i+1}), \ell(\lambda_{i+1}) \leq m$ such that $\mu_{i+1} u'_i \lambda_{i+1} = u'_{i+1}$ in H. The details of the proof of this depend on which of cases (1) and (2) applies and how u'_i is positioned in relation to the various subwords. For instance suppose we are in the first case, so that $u_i = \alpha_i s^{\pm1} x_1 \cdots x_n s^{\mp1} \beta_i$ and $u_{i+1} = \alpha_i \varphi^{\mp1}(x_1) \cdots \varphi^{\mp1}(x_n) \beta_i$ and suppose $u'_{i+1} = \alpha'_i \varphi^{\mp1}(x_1) \cdots \varphi^{\mp1}(x_j) \gamma$ where α'_i is a suffix of α_i and γ is a prefix of $\varphi^{\mp1}(x_{j+1})$. Then taking $u'_i = \alpha'_i s^{\pm1} x_1 \cdots x_j$, the result holds with μ_{i+1} the empty word and $\lambda_{i+1} = s^{\mp1}\gamma$. (The length of γ is strictly less than $\ell(\varphi^{\mp1}(x_{j+1}))$, which is at most m—see equation (2).) The other cases are similar.

Take u'_r to be a subword of u_r which freely reduces to u'. As per the previous paragraph obtain $u'_{r-1}, \ldots, u'_0$ and $\mu_r, \ldots, \mu_1$ and $\lambda_r, \ldots, \lambda_1$ such that $\mu_r \cdots \mu_1 u'_0 \lambda_1 \cdots \lambda_r = u'_r = u'$ in H. But $\ell(\mu_r \cdots \mu_1 u'_0 \lambda_1 \cdots \lambda_r) \leq (2m+1)\ell(u)$ since u'_0 is a subword of u, $r \leq \ell(u)$ and $\ell(\mu_i), \ell(\lambda_i) \leq m$ for all i.

So when u is a minimal length word representing h in H, we get our result. □

We introduce some notation. If w is the reduced word representing $h \in F$, then $i = \operatorname{rank}(w)$ is the maximum i such that there is a letter $a_i^{\pm1}$ in w, and then $|h|_\pi$ denotes the number of pieces in the rank-i piece decomposition of w. Further, for a word w and a letter a, the number of occurrences of a in w plus the number of a^{-1} is $\mathrm{wt}_a(w)$, and the number of occurrences of a minus the number of a^{-1} is $\exp_a(w)$.

Lemma 4.8. *Suppose $h \in H$ is expressed in normal form as $\tilde{u}s^r$, where $\tilde{u}$ is a reduced word on $a_1^{\pm1}, \ldots, a_m^{\pm1}$ and $r \in \mathbb{Z}$. Then $|\tilde{u}|_\pi \leq |h|_H$.*

Proof. Let v_m be a geodesic word on $a_1^{\pm1}, \ldots, a_m^{\pm1}, s^{\pm1}$ representing h in H. So $\ell(v_m) = |h|_H$.

The *shuffling* moves $sa_i \mapsto \varphi^{-1}(a_i)s$ and $s^{-1}a_i \mapsto \varphi(a_i)s^{-1}$ transform v_m to a word $u_m s^r$ so that $\tilde{u}$ is the freely reduced form of u_m. Since these moves do not create or remove them, the $a_m^{\pm1}$ in u_m correspond with those in v_m in number, sign, and relative location.

Say that a subword σ in v_m is *superfluous* if it has the form $a_m \tau a_m^{-1}$ or $a_m^{-1}\tau a_m$ for some word τ on $a_1^{\pm1}, \ldots, a_m^{\pm1}$ and the $a_m^{\pm1}$ and $a_m^{\mp1}$ that bookend σ, correspond to an $a_m^{\pm1}$ and an $a_m^{\mp1}$ that bookend a subword $\overline{\sigma}$ in u_m that freely reduces to the identity. Given such a σ, write $v_m = \sigma_0\sigma\sigma_1$ (as words), let $\lambda = \exp_s(\tau)$, and let $\mu = \exp_s(\sigma_0)$. There exists a word ξ on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$ arising in the following three cases as follows:

(1) If $\sigma = a_m \tau a_m^{-1}$, then the shuffling moves give that $\sigma = a_m \xi a_m^{-1} s^{\lambda}$ in H. Further, $\varphi^{-\mu}(a_m \xi a_m^{-1}) = \overline{\sigma}$. Define $\sigma' = s^{\lambda}$.
(2) If $\sigma = a_m^{-1} \tau a_m$ and $\lambda \leq 0$, then they give that $\sigma = a_m^{-1} \xi a_m a_{m-1}^{-\lambda} s^{\lambda}$ in H. Further, $\varphi^{-\mu}(\xi) = \overline{\sigma}$. Define $\sigma' = a_{m-1}^{-\lambda} s^{\lambda}$.
(3) If $\sigma = a_m^{-1} \tau a_m$ and $\lambda > 0$, then moves $a_i s \mapsto s\varphi(a_i)$ and $a_i s^{-1} \mapsto s^{-1}\varphi^{-1}(a_i)$ give that $\sigma = s^{\lambda} a_{m-1}^{-\lambda} a_m^{-1} \xi a_m$ in H. Further, $\varphi^{-\mu-\lambda}(\xi) = \overline{\sigma}$. Define $\sigma' = s^{\lambda} a_{m-1}^{-\lambda}$.

In each case $\xi = 1$, because $\overline{\sigma} = 1$, and so $\sigma = \sigma'$ in H.

Let $i = \operatorname{rank}(\tilde{u})$. If $i < m$, then all the $a_m^{\pm 1}$ in u_m cancel away on free reduction of u_m. So there exists a family of pairwise disjoint superfluous subwords of v_m which together contain every $a_m^{\pm 1}$ in v_m. Let v_{m-1} be the word obtained from v_m by replacing each of these superfluous subwords σ by the corresponding σ' described above. Then $v_{m-1} = v_m$ in H and

$$\begin{aligned} \operatorname{rank}(v_{m-1}) &\leq m-1, \\ \operatorname{wt}_s(v_{m-1}) &\leq \operatorname{wt}_s(v_m), \\ \operatorname{wt}_{a_j}(v_{m-1}) &\leq \operatorname{wt}_{a_j}(v_m) \text{ for } j = 1, \ldots, m-2, \\ \operatorname{wt}_{a_{m-1}}(v_{m-1}) &\leq \operatorname{wt}_{a_{m-1}}(v_m) + \operatorname{wt}_s(v_m), \end{aligned}$$

where the final inequality holds because the $a_{m-1}^{-\lambda}$ inserted in all instances of cases (2) and (3) contribute a total of no more than $\operatorname{wt}_s(v_m)$ letters $a_{m-1}^{\pm 1}$.

If $i < m-1$, then, because there are no $a_m^{\pm 1}$ letters in v_{m-1}, we can obtain a word v_{m-2} from v_{m-1} in the same manner that we obtained v_{m-1} from v_m and subject to the same inequalities as displayed above, but with m decremented by 1. Repeat until arriving at v_i. Then $\operatorname{rank}(v_i) = i$ and

$$\operatorname{wt}_{a_1}(v_i) + \cdots + \operatorname{wt}_{a_i}(v_i) \leq \operatorname{wt}_{a_1}(v_m) + \cdots + \operatorname{wt}_{a_i}(v_m) + \operatorname{wt}_s(v_m) = \ell(v_m). \tag{4}$$

Now, v_i freely equals $w_0 s^{\alpha_1} w_1 s^{\alpha_2} w_2 \cdots s^{\alpha_k} w_k$ for some reduced words $w_0, \ldots, w_k$ on $a_1^{\pm 1}, \ldots, a_i^{\pm 1}$ and some non-zero $\alpha_1, \ldots, \alpha_k \in \mathbb{Z}$. For $0 \leq j \leq k$, let $\beta_j = -\alpha_1 - \cdots - \alpha_j$, so that $\tilde{u}$ freely equals $\varphi^{\beta_0}(w_0) \cdots \varphi^{\beta_k}(w_k)$. The number of pieces in the rank-i decomposition of w_j is at most $\ell(w_j)$ since each piece has at least one letter. By Lemma 3.2, the rank-i decompositions of w_j and of $\varphi^{\beta_j}(w_j)$ have the same number of pieces. Free reduction between an $\varphi^{\beta_j}(u_j)$ and the neighbouring $\varphi^{\beta_{j+1}}(u_{j+1})$ can only cause pieces to merge or cancel, so $|\tilde{u}|_\pi \leq \ell(w_0) + \cdots + \ell(w_k)$, which is at most $\ell(v_m)$ by (4). □

Corollary 4.9. *For $i = 1, \ldots, m$, there exists $K_i > 0$ such that for all $g \in F$ of rank i, we have $|g|_F \leq K_i |g|_\pi \, |g|_H^{i-1}$.*

Proof. We induct on i. If $i = 1$, then $g = a_1^k$ for some $k \in \mathbb{Z}$, and $|k| = |g|_F = |g|_\pi$.

Now assume $i > 1$. Express g, viewed as a reduced word on $a_1^{\pm 1}, \ldots, a_i^{\pm 1}$, as a rank-$i$ product of pieces $\pi_1 \cdots \pi_p$. Each piece is $\pi_k = a_i^{\delta_k} v_k a_i^{-\varepsilon_k}$ for some $\delta_k, \varepsilon_k \in \{0, 1\}$ and v_k a reduced word of rank at most $i-1$. By induction $|v_k|_F \leq K_{i-1} |v_k|_\pi \, |v_k|_H^{i-2}$. By Lemma 4.8 $|v_k|_\pi \leq |v_k|_H$. So $|v_k|_F \leq K_{i-1} \, |v_k|_H^{i-1}$. We then get

$$|g|_F = \sum_{k=1}^{p} |\pi_k|_F \leq \sum_{k=1}^{p} (K_{i-1} \, |v_k|_H^{i-1} + 2) \leq |g|_\pi (K_{i-1} (2m+1)^{i-1} \, |g|_H^{i-1} + 2)$$

where the last inequality follows from Proposition 4.5. □

If $g = s^{-k}a_i^k s^k = \varphi^k(a_i^k)$, then $|g|_H \leq 3k$ and g is a product of k pieces, each of which is $\varphi^k(a_i)$, a positive word whose length grows like a polynomial in k of degree $i-1$, as we saw in Lemma 4.2. Thus the bound in Corollary 4.9 is sharp in this case (up to constants).

5. Solving the 0-twisted conjugacy problem

As we saw in Section 2, the conjugacy relation $uw = wv$ in H amounts to the φ-twisted conjugacy relation $\tilde{u}\varphi^{-p}(\tilde{w}) = \tilde{w}\varphi^{-r}(\tilde{v})$ in F, where we have normal forms $u = \tilde{u}s^p$, $v = \tilde{v}s^p$, and $w = \tilde{w}s^r$. We assume in this section that $p = 0$.

Recall that the 0-twisted conjugacy problem asks whether, given words $\tilde{u}, \tilde{v}$ on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$, there exist $r \in \mathbb{Z}$ and $\tilde{w} \in F$ such that

$$\tilde{u}\tilde{w} \; = \; \tilde{w}\varphi^{-r}(\tilde{v}) \;\text{ in } F. \tag{5}$$

Proposition 5.1 (0-twisted conjugacy problem)**.** *There exists $A > 0$ with the following property. Suppose $\tilde{u}$ and $\tilde{v}$ are words on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$.*

(I) If there exist $r \in \mathbb{Z}$ and $\tilde{w} \in F$ satisfying (5), then there are such r and $\tilde{w}$ with $|r| + |\tilde{w}|_H \leq A(|\tilde{u}|_H + |\tilde{v}|_H)$.
(II) If there exists $\tilde{w} \in F$ satisfying (5) with $r = 0$, then there exists such $\tilde{w}$ with $|\tilde{w}|_H \leq A(|\tilde{u}|_H + |\tilde{v}|_H)$.
(III) There is an algorithm that, given $\tilde{u}$ and $\tilde{v}$ will determine whether or not there exist $r \in \mathbb{Z}$ and $\tilde{w} \in F$ solving (5), and will exhibit them if they exist. The running time of this algorithm is polynomial in $|\tilde{u}|_H + |\tilde{v}|_H$.

The proof of Proposition 5.1 uses the following lemma, which determines the form of a 'short' solution to (5) whenever any solution exists.

Lemma 5.2. *There exists $B > 0$ with the following property. Suppose $\tilde{u}, \tilde{v} \in F$ are as in Proposition 5.1. Suppose there exist $r \in \mathbb{Z}$ and $\tilde{w} \in F$ satisfying (5). Then there exist $\tilde{w}_0 \in F$ satisfying*

$$\tilde{u}\tilde{w}_0 \; = \; \tilde{w}_0\varphi^{-r}(\tilde{v}) \;\text{ in } F, \tag{6}$$

with $\tilde{w}_0 = UV$, where U is a prefix of $\tilde{u}^{-1}$ and $\varphi^r(V)$ is a suffix of $\tilde{v}$.

Furthermore, either $\tilde{u}\tilde{w} = \tilde{w}\tilde{v}$ in F, or $|r| \leq B(|\tilde{u}|_H + |\tilde{v}|_H)$.

Proof. First replace $\tilde{u}$ by a cyclic conjugate u' such that $u'u'$ is reduced (i.e. u' is cyclically reduced) and the rank-m piece decomposition of $u'\,u'$ is the concatenation of two copies of the piece decomposition of u'—that is, the rightmost piece in u' does not combine with the leftmost piece in u' to make a single piece in $u'\,u'$. This is achieved by conjugating $\tilde{u}$ by a suitable y that is a prefix of $\tilde{u}^{-1}$, so that $y^{-1}\tilde{u}y = u'$ in F.

Likewise, replace $\tilde{v}$ by a similarly structured cyclic conjugate v'. Let z be the prefix of $\tilde{v}^{-1}$ such that $z^{-1}\tilde{v}z = v'$ in F.

Lemma 3.2 gives us that $\varphi^k(u')$ and $\varphi^k(v')$ are cyclically reduced for all $k \in \mathbb{Z}$.

By assumption, we have $r \in \mathbb{Z}$ and $\tilde{w} \in F$ satisfying (5). This implies that there is $\tilde{w}_1 \in F$ satisfying

$$u' \tilde{w}_1 \ = \ \tilde{w}_1 \varphi^{-r}(v') \quad \text{in } F. \tag{7}$$

Since (7) is a conjugacy relation in F and $\varphi^{-r}(v')$ is cyclically reduced, Lemma 2.1 tells us that there is a prefix $\tilde{w}_2$ of $(u')^{-1}$ such that $u'\tilde{w}_2 = \tilde{w}_2\varphi^{-r}(v')$ in F. Then $\tilde{w}_0 = y\tilde{w}_2\varphi^{-r}(z^{-1})$ satisfies (6) with $r_0 = r$. We then take $U = y\tilde{w}_2$, which is a prefix of $\tilde{u}^{-1}$ (since y is, and $\tilde{w}_2$ is a prefix of $(u')^{-1} = y^{-1}\tilde{u}^{-1}y$) and $V = \varphi^{-r}(z^{-1})$. From the definition of z, $\varphi^{r}(V) = z^{-1}$ is a suffix of $\tilde{v}$.

To complete the proof, we need to bound $|r|$.

If v' is fixed by φ or $r = 0$, then v' is conjugate to u' in F and we have $\tilde{u}\tilde{w} = \tilde{w}\tilde{v}$ in F.

Suppose, then, that v' is not fixed by φ and $r \neq 0$. Let i be the rank of v'. As v' is not fixed by φ, Lemma 4.1 tells us that it has a piece π that is itself of rank i and is not fixed by φ. Proposition 4.4 and Corollary 4.9 give constants $C_i, K_i > 0$ such that either

$$C_i\, |r|^{i-1} \ < \ |\pi|_F \ \leq \ K_i\, |\pi|_H^{i-1}\,, \tag{8}$$

or

$$C_i \left(|r| - \left(\frac{|\pi|_F}{C_i} \right)^{\frac{1}{i-1}} \right)^{i-1} \ \leq \ \left|\varphi^{-r}(\pi)\right|_F \ \leq \ K_i \left|\varphi^{-r}(\pi)\right|_H^{i-1}. \tag{9}$$

Since π is a subword of v', which is a subword of $\tilde{v}$, Proposition 4.5 gives us $|\pi|_H \leq (2m+1)\,|\tilde{v}|_H$. Then by Corollary 4.9 we get $|\pi|_F \leq K_i(2m+1)^{i-1}\,|\tilde{v}|_H^{i-1}$. Meanwhile, since u' and $\varphi^{-r}(v')$ are cyclically reduced and conjugate in F, it follows that $\varphi^{-r}(\pi)$ is a subword of $u'u'$. Since the piece decomposition of $u'u'$ consists of the concatenation of two copies of that of u', we must have that $\varphi^{-r}(\pi)$ is a subword of u'. So $|\varphi^{-r}(\pi)|_H \leq (2m+1)\,|\tilde{u}|_H$ by Proposition 4.5. Both (8) and (9) lead to

$$|r| \ \leq \ \left(\frac{K_i}{C_i} \right)^{\frac{1}{i-1}} (2m+1)(|\tilde{u}|_H + |\tilde{v}|_H),$$

showing that a suitable $B > 0$ exists. □

Proof of Proposition 5.1. We begin by establishing the length bounds in (I) and (II).

Since in Lemma 5.2 the value of r does not change between (5) and (6), case (II) of Proposition 5.1 holds. Indeed, we need only use Proposition 4.5 to bound $|U|$ and $|V|$. Similarly, the bound on $|r|$ and the form of $\tilde{w}_0$ give (I).

Next we consider the complexity of the algorithms solving the 0-twisted conjugacy problem and its search variant.

Lemma 5.2 tells us that if a solution exists, then there is a solution of a particularly nice form. On input $\tilde{u}$ and $\tilde{v}$, we list all pairs $(\tilde{w}, r)$, where r is an integer satisfying $|r| \leq B(|\tilde{u}|_H + |\tilde{v}|_H)$, and $\tilde{w}$ has the form UV, with U a prefix of $\tilde{u}^{-1}$ and $V = \varphi^{-r}(\hat{V})$, where $\hat{V}$ is a suffix of $\tilde{v}$. It is not hard to see that the number of such pairs is polynomially bounded in terms of $|u|_H + |v|_H$. So a search through this list for a solution to (5) can be completed in polynomial time. If none is found, we conclude that no solution exists. □

6. Preserved prefixes

Recall that $uw = wv$ in H amounts to $\tilde{u}\varphi^{-p}(\tilde{w}) = \tilde{w}\varphi^{-r}(\tilde{v})$ in F, where $u = \tilde{u}s^p$, $v = \tilde{v}s^p$, and $w = \tilde{w}s^r$ are the normal forms. We now assume $p \neq 0$. In the specific case when $\tilde{u}$ is the empty word, $\tilde{w}$ is a concatenation of a prefix of $\varphi^{-p}(\tilde{w})$ with a subword of $\varphi^{-r}(\tilde{v})^{-1}$. This points to the fact that, in order to understand the length (or structure) of $\tilde{w}$ (and hence w), it is important to understand the length (or structure) of the longest common prefix of $\tilde{w}$ and $\varphi^{-p}(\tilde{w})$.

We begin with the instance where $\tilde{w}$ is a single piece of a type that behaves well with regards to common prefixes. This will feed into the general case in Corollary 6.2 below.

If π is $a_2a_1^q$ for some $q \geq 0$, then (assuming $r \geq 0$) the length $1 + q$ of the longest common prefix $a_2a_1^q$ of π and $\varphi^r(\pi) = a_2a_1^{q+r}$ can be arbitrarily large compared to $\left|\pi^{-1}\varphi^r(\pi)\right|_H = \left|a_1^r\right|_H = r$. The same can be said when π is a_1^q or $a_1^qa_i^{-1}$ for any $i \geq 2$ and $q \in \mathbb{Z}$. In contrast, for other types of pieces, the form of the longest common prefix is constrained in a manner that strongly restricts its length:

Lemma 6.1. *For $3 \leq i \leq m$, there exists $B_i > 0$ with the following property. Suppose $r > 0$ and that π is a rank-i piece whose first letter is a_i. Then the longest common prefix L of π and $\varphi^r(\pi)$ is a concatenation $L = \Lambda_1\Lambda_2$ of words, where*

- *Λ_1 is a prefix of $\varphi^k(a_i)$ for some $k \in \mathbb{Z}$ satisfying $|k| \leq B_i\left(\left|\pi^{-1}\varphi^r(\pi)\right|_H + |r|\right)$,*
- *Λ_2 is a subword of $\varphi^r(a_i^{-1})$.*

Proof. We may assume that $\ell(L) \geq 2$, else $L = a_i$ and the result is immediate with $L = \Lambda_1$ and $k = 0$.

Case 1. $\pi = a_iu$ for a word u of rank less than i.

We claim that either there is no cancellation (as in Figure 3) between $\varphi^r(a_i)$ and $\varphi^r(u)$, or there is *complete cancellation* (as in Figure 2) by which we mean that $\varphi^r(a_i)\varphi^r(u)$, freely reduces to a_i times a suffix of $\varphi^r(u)$. After all, if there is not complete cancellation, then the first two letters a_ia_{i-1} of $\varphi^r(a_i)$ are not cancelled away on free reduction of $\varphi^r(a_i)\varphi^r(u)$. As $\ell(L) \geq 2$, these are also the first two letters of $\pi = a_iu$, and so the first letter of u is a_{i-1}. But then the first letter of $\varphi^r(u)$ must also be a_{i-1}, and as $\varphi^r(a_i)$ is a positive word, there is no cancellation between it and $\varphi^r(u)$.

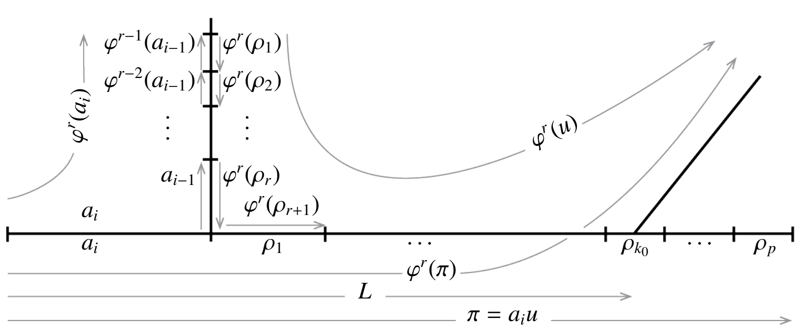

Figure 2. Cancellation as per Case 1a of the proof of Lemma 6.1.

Let $\rho_1 \cdots \rho_p$ be the rank-$(i-1)$ decomposition of u into pieces. By Lemma 3.2, $\varphi^r(\rho_1)\cdots\varphi^r(\rho_p)$ is the rank-$(i-1)$ piece decomposition of $\varphi^r(u)$. Choose k_0 so that

(10) $$L \;=\; a_i\rho_1\cdots\rho_{k_0-1}\rho'_{k_0},$$

where k_0 is chosen so that ρ'_{k_0} is a non-empty prefix of ρ_{k_0}.

Case 1a. Complete cancellation.

We will show that $a_i\rho_1\cdots\rho_{k_0} = \varphi^{-k_0}(a_i)$, and so L is a prefix of this and the result will hold with $L = \Lambda_1$ and Λ_2 the empty word.

Lemma 4.3 tells us that $\varphi^r(a_i) = a_i\, a_{i-1}\varphi(a_{i-1})\cdots\varphi^{r-1}(a_{i-1})$. From this we can read off the first r pieces of $\varphi^r(u)$ on account of the 'complete cancellation' between $\varphi^r(a_i)$ and $\varphi^r(u)$: for $k = 1,\ldots,r$ we have $\varphi^r(\rho_k) = \varphi^{r-k}(a_{i-1})^{-1}$, or equivalently

(11) $$\rho_k \;=\; \varphi^{-k}(a_{i-1})^{-1}.$$

After 'complete cancellation' $\varphi^r(\pi) = a_i\varphi^r(\rho_{r+1})\cdots\varphi^r(\rho_p)$. As L is also a prefix of $\varphi^r(\pi)$,

$$L \;=\; a_i\varphi^r(\rho_{r+1})\cdots\varphi^r(\rho_{r+k_0-1})\rho'_{k_0}$$

where, comparing with (10), for $k = 1,\ldots,k_0-1$ we have $\rho_k = \varphi^r(\rho_{k+r})$, or equivalently $\rho_{k+r} = \varphi^{-r}(\rho_k)$. By induction, we can extend (11) to $k = 1,\ldots,k_0{+}r-1$. This tells us in particular that

(12) $$\rho_{k_0} \;=\; \varphi^{-k_0}(a_{i-1})^{-1},$$

since $r > 0$, and

$$a_i\rho_1\cdots\rho_{k_0} \;=\; a_i\varphi^{-1}(a_{i-1})^{-1}\cdots\varphi^{-k_0}(a_{i-1})^{-1},$$

which equals, as a word, $\varphi^{-k_0}(a_i)$ by Lemma 4.3. It follows that L is a prefix of $\varphi^{-k_0}(a_i)$.

Next we will give an upper bound on k_0 that will imply an upper bound on $|-k_0+1|$, proving the condition on Λ_1. From (12) we get

(13) $$C_i k_0^{i-1} \;\le\; \left|\rho_{k_0}\right|_F$$

by Lemma 4.2. As ρ_{k_0} is a single piece, Corollary 4.9 gives

(14) $$\left|\rho_{k_0}\right|_F \;\le\; K_i\left|\rho_{k_0}\right|_H^{i-1}.$$

View ρ_{k_0} as a product of ρ'_{k_0} with a subword of $\pi^{-1}\varphi^r(\pi)$. Then apply Proposition 4.5 to give

(15) $$\left|\rho_{k_0}\right|_H \;\le\; \left|\rho'_{k_0}\right|_H + (2m+1)\left|\pi^{-1}\varphi^r(\pi)\right|_H.$$

To bound $\left|\rho'_{k_0}\right|_H$, observe that ρ'_{k_0} is a prefix of $\varphi^r(\rho_{k_0+r})$, and ρ_{k_0+r} is a subword of $\pi^{-1}\varphi^r(\pi)$, since $r > 0$. So, applying Proposition 4.5, we first get

$$\left|\rho'_{k_0}\right|_H \;\le\; (2m+1)\left|\varphi^r(\rho_{k_0+r})\right|_H.$$

Then using that $\varphi^r(\rho_{k_0+r}) = s^{-r}\rho_{k_0+r}s^r$, we deduce that

$$\left|\rho'_{k_0}\right|_H \;\le\; (2m+1)\left(2r + \left|\rho_{k_0+r}\right|_H\right).$$

A last application of Proposition 4.5 then gives

(16) $$\left|\rho'_{k_0}\right|_H \;\le\; (2m+1)\left(2r + (2m+1)\left|\pi^{-1}\varphi^r(\pi)\right|_H\right).$$

Together (13)–(16) show k_0 is at most a constant times $|r| + \left|\pi^{-1}\varphi^r(\pi)\right|_H$, as required.

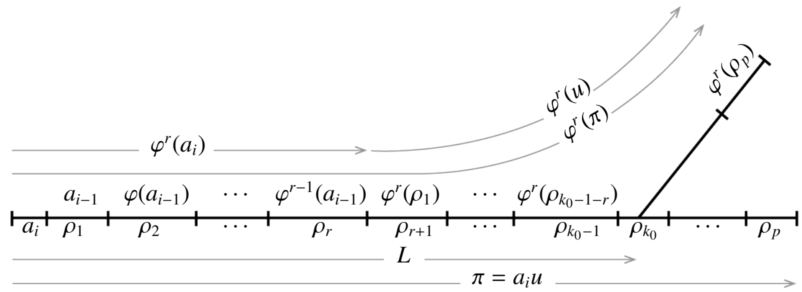

FIGURE 3. Cancellation as per Case 1b of the proof of Lemma 6.1.

Case 1b. No cancellation.

We will show that L is a prefix of $\varphi^{k_0}(a_i)$ and that the result will again hold with $L = \Lambda_1$ and Λ_2 the empty word.

Comparing pieces along the common prefix $L = a_i\rho_1 \cdots \rho_{k_0-1}\rho'_{k_0}$ of $\pi = a_i\rho_1 \cdots \rho_p$ and

$$\varphi^r(\pi) \;=\; \varphi^r(a_i)\varphi^r(\rho_1)\cdots\varphi^r(\rho_p) \;=\; a_i\, a_{i-1}\varphi(a_{i-1})\cdots\varphi^{r-1}(a_{i-1})\varphi^r(\rho_1)\cdots\varphi^r(\rho_p),$$

which is a freely reduced word in this case, we claim that

$$\rho_k \;=\; \varphi^{k-1}(a_{i-1}) \;\text{ for }\; k = 1,\dots,k_0-1. \tag{17}$$

For $k \le \min\{r, k_0-1\}$ we get $\rho_k = \varphi^{k-1}(a_{i-1})$ immediately, so there is nothing left to show when $r \ge k_0 - 1$. When $r < k_0 - 1$, if $k = r+1,\dots,k_0-1$ then $\rho_k = \varphi^r(\rho_{k-r})$, and this gives the claim inductively.

We claim that ρ'_{k_0} is a prefix of $\varphi^{k_0-1}(a_{i-1})$. Indeed, if $k_0 \le r$, then this is immediate. Meanwhile, if $k_0 > r$, ρ'_{k_0} is a subword of $\varphi^r(\rho_{k_0-r})$ which equals $\varphi^{k_0-1}(a_{i-1})$ by (17) since $r > 0$.

It follows that L is a prefix of $a_ia_{i-1}\varphi(a_{i-1})\cdots\varphi^{k_0-1}(a_{i-1})$, which equals, as a word, $\varphi^{k_0}(a_i)$.

We complete this case by bounding k_0. The process is similar to Case 1a.

Since $\varphi^r(\rho_{k_0})$ is a subword of $\pi^{-1}\varphi^r(\pi)$, we can obtain a bound on the length of ρ'_{k_0} as follows. Firstly, $\left|\rho'_{k_0}\right|_H \le (2m+1)\left|\rho_{k_0}\right|_H$ by Proposition 4.5. Then, using $\rho_{k_0} = s^r\varphi^r(\rho_{k_0})s^{-r}$, and that $\varphi^r(\rho_{k_0})$ is a subword of $\pi^{-1}\varphi^r(\pi)$, we get $\left|\rho_{k_o}\right|_H \le 2\,|r| + (2m+1)\left|\pi^{-1}\varphi^r(\pi)\right|_H$. Thus, we have

$$\left|\rho'_{k_0}\right|_H \;\le\; (2m+1)\left(2\,|r| + (2m+1)\left|\pi^{-1}\varphi^r(\pi)\right|_H\right).$$

Since $\varphi^r(\rho_{k_0-r}) = \varphi^{k_0-1}(a_{i-1})$, Lemma 4.2 and Corollary 4.9 imply that

$$C_i(k_0-1)^{i-1} \;\le\; \left|\varphi^r(\rho_{k_0-r})\right|_F \;\le\; K_i\left|\varphi^r(\rho_{k_0-r})\right|_H^{i-1}.$$

We can write $\varphi^r(\rho_{k_0-r})$ as a product of ρ'_{k_0} and a subword of $\pi^{-1}\varphi^r(\pi)$. Hence, by Proposition 4.5,

$$\left|\varphi^r(\rho_{k_0-r})\right|_H \;\le\; \left|\rho'_{k_0}\right|_H + (2m+1)\left|\pi^{-1}\varphi^r(\pi)\right|_H.$$

These displayed inequalities combine to give an upper bound on k_0 implying the condition on Λ_1.

Case 2. $\pi = a_i u a_i^{-1}$ for a reduced word u of rank less than i.

Let $\pi_0 = a_i u$. The only a_i^{-1} in π is at the end; ditto in $\varphi^r(\pi)$. So if the common prefix L of π and $\varphi^r(\pi)$ is the whole of π, then $L = \varphi^r(\pi)$, also, but that cannot be: by Lemma 4.1 π is not fixed by φ and then by Proposition 4.4 it is not fixed by φ^r (as $r \neq 0$). So L is, in fact, a prefix of π_0.

As $\varphi^r(\pi)$ is the free reduction of $\varphi^r(\pi_0)\varphi^r(a_i^{-1})$, the word L is the concatenation $\Lambda_1\Lambda_2$ of a prefix Λ_1 of $\varphi^r(\pi_0)$ with a subword Λ_2 of $\varphi^r(a_i^{-1})$. But then Λ_1 is a common prefix of π_0 and $\varphi^r(\pi_0)$ (though it may not be the full common prefix) so we deduce from Case 1 that Λ_1 is a prefix of $\varphi^k(a_i)$, where $|k|$ is bounded by a constant times $r + \left|\pi_0^{-1}\varphi^r(\pi_0)\right|_H$. Since $\pi_0^{-1}\varphi^r(\pi_0) = a_i^{-1}\pi^{-1}\varphi^r(\pi)a_i$, we have $\left|\pi_0^{-1}\varphi^r(\pi_0)\right|_H \leq \left|\pi^{-1}\varphi^r(\pi)\right|_H + 2$ and the required bound on $|k|$ follows. $\square$

Corollary 6.2. *For all $1 \leq i \leq m$, there exists $A_i > 0$ with the following property. For all freely reduced $w \in F$ of rank i and all $r \in \mathbb{Z}$, there exists a freely reduced word $w_0 \in F$ of rank at most i such that the following hold.*

- ($\mathcal{P}1$) *If the free reduction of $w^{-1}\varphi^r(w)$ is $\alpha\beta$, with α a prefix of w^{-1} and β a suffix of $\varphi^r(w)$, then the free reduction of $w_0^{-1}\varphi^r(w_0)$ is $\alpha\beta$, and α a prefix of w_0^{-1} and β a suffix of $\varphi^r(w_0)$.*
- ($\mathcal{P}2$) *The longest common prefix P of w_0 and $\varphi^r(w_0)$ has the form $P = P_1P_3\cdots P_i$ where*
  - *P_1 is a prefix of $\varphi^k(a_t)$ for some $t \leq i$ and some $k \in \mathbb{Z}$ satisfying $|k| \leq A_i\,(|\alpha\beta|_H + |r|)$, and*
  - *P_j is a subword of $\varphi^r(a_j^{-1})$ for $j = 3, \ldots, i$.*

Proof. If $w = \varphi^r(w)$, then we can take α, β, and w_0 to be the empty word. So assume $w \neq \varphi^r(w)$. In particular, $r \neq 0$.

The statement for $r < 0$ will follow from that for $r > 0$ since we could instead consider the common prefix of $\overline{w} = \varphi^r(w)$ and $\varphi^{-r}(\overline{w})$, which of course equals P. So assume $r > 0$.

We will induct on $i = \operatorname{rank}(w)$. The case $i = 1$ is elementary: $w = \varphi^r(w)$ and the result holds as we just explained.

Assume $i > 1$. Let $w = \pi_1 \cdots \pi_p$ be the rank-i decomposition of w into pieces. Our first step is to reduce the problem to a question concerning a single piece. Take k minimal so that the longest common prefix of w and $\varphi(w)$ is a subword of $\pi_1 \cdots \pi_k$. Let $\pi := \pi_k$. It follows from Lemma 3.2 that $\pi_1, \ldots, \pi_{k-1}$ are fixed by φ^r. So ($\mathcal{P}1$) amounts to:

- ($\mathcal{P}1'$) If the free reduction of $w^{-1}\varphi^r(w)$ is $\alpha\beta$, with α a prefix of w^{-1} and β a suffix of $\varphi^r(w)$, then the free reduction of $(\pi_k\cdots\pi_p)^{-1}\varphi^r(\pi_k\cdots\pi_p)$ is $\alpha\beta$, with α a prefix of $(\pi_k\cdots\pi_p)^{-1}\varphi^r(\pi_k\cdots\pi_p)$ and β' a suffix of $\varphi^r(\pi)$.

We will find a word π_0 such that the longest common prefix P of π_0 and $\varphi^r(\pi_0)$ satisfies the conditions for ($\mathcal{P}2$), and π_0 satisfies:

- ($\mathcal{P}1''$) If the free reduction of $\pi^{-1}\varphi^r(\pi)$ is $\alpha'\beta'$, with α' a prefix of π^{-1} and β' a suffix of $\varphi^r(\pi)$, then the free reduction of $\pi_0^{-1}\varphi^r(\pi_0)$ is $\alpha'\beta'$, with α' a prefix of π_0^{-1} and β' a suffix of $\varphi^r(\pi_0)$.

Then, setting $w_0 = \pi_0\pi_{k+1}\dots\pi_p$, we will get $\alpha\beta = w_0^{-1}\varphi^r(w_0)$ in F by ($\mathcal{P}1'$), with α a prefix of w_0^{-1} and β a suffix of $\varphi^r(w_0)$, satisfying ($\mathcal{P}1$). We claim that P being the longest common prefix of π_0 and $\varphi^r(\pi_0)$ implies it is also the longest common prefix of w_0 and $\varphi^r(w_0)$. Indeed, if $w_0 = \pi_0$ (as words), this is immediate. And if $w_0 \neq \pi_0$, then either the prefixes are as claimed, or $\pi_0 = \varphi^r(\pi_0)$, implying $\pi^{-1}\varphi^r(\pi)$ is trivial, by ($\mathcal{P}1''$), contradicting our choice of k.

Now we turn to finding this π_0.

Assume that $i = 2$, which is an exceptional case. By the minimality of k, we cannot have π equal to a_1^q or $a_2a_1^qa_2^{-1}$ for some $q \in \mathbb{Z}$, for that would imply that $\pi = \varphi^r(\pi)$. The remaining possibilities are that π is $a_2a_1^q$ or $a_1^qa_2^{-1}$ for some $q \in \mathbb{Z}$. Take $\pi_0 = a_2$ or $\pi_0 = a_2^{-1}$, respectively in these two cases. In the former case, α' is the empty word, and $\beta' = a_1^r$. In the latter, $\alpha' = a_2$ and $\beta' = a_1^{-r}a_2^{-1}$. Both satisfy ($\mathcal{P}1''$). The longest common prefix of π_0 and $\varphi^r(\pi_0)$, and hence also of w_0 and $\varphi^r(w_0)$, is then either a_2 or the empty word. Taking P_1 to be a_2 or the empty word, accordingly, and $P_3,\dots,P_i$ all the empty word gives us the required form ($\mathcal{P}2$).

Suppose $i \geq 3$. How we proceed depends on the type and rank of the piece π.

Case 1. $j := \operatorname{rank}(\pi) < i$.

Apply the induction hypothesis to get a freely reduced word π_0 of rank at most j satisfying ($\mathcal{P}1''$) and such that the longest common prefix of π_0 and $\varphi^r(\pi_0)$ has the form $P = P_1P_3\cdots P_j$, where $P_3,\dots,P_j$ are each subwords of $\varphi^r(a_j^{-1})$, and P_1 a prefix of $\varphi^k(a_t)$, for some $t \leq j$ and some $|k| \leq A_j(\left|\pi^{-1}\varphi^r(\pi)\right|_H + |r|)$. Since $\pi^{-1}\varphi^r(\pi)$ is a subword of $\alpha\beta$, taking A_i large enough that $A_i \geq (2m+1)A_j$ will mean, by Proposition 4.5, that P_1 satisfies the requirements stated in this corollary. Finally, we take $P_{j+1},\dots,P_i$ to all be the empty word.

Case 2. The first letter of π is a_i.

In this case we take $\pi_0 = \pi$, so ($\mathcal{P}1''$) trivially holds. Lemma 6.1 gives us the structure of P as required, with $P_1 = \Lambda_1$, with $P_3,\dots,P_{i-1}$ being empty words, and with $P_i = \Lambda_2$. We just note that the power k in P_1 satisfies the required bound by taking A_i large enough so that $A_i \geq (2m+1)B_i$, where B_i is the constant from Lemma 6.1 (as in Case 1, this is because $\pi^{-1}\varphi^r(\pi)$ is a subword of $\alpha\beta$, and we can apply Proposition 4.5).

Case 3. $\pi = ua_i^{-1}$ with $j := \operatorname{rank}(u) < i$.

By the inductive hypothesis, there is a word $u_0 \in F$ of rank at most j such that

($\mathcal{P}1'''$) If the free reduction of $u^{-1}\varphi^r(u)$ is $\gamma\delta$, with γ a prefix of u^{-1} and δ a suffix of $\varphi^r(u)$, then the free reduction of $u_0^{-1}\varphi^r(u_0)$ is $\gamma\delta$, with γ a prefix of u_0^{-1} and δ a suffix of $\varphi^r(u_0)$,

and the maximal common prefix of u_0 and $\varphi^r(u_0)$ is of the form $P_0 = P_1P_3\cdots P_j$, with $P_1, P_3\dots,P_j$ as per ($\mathcal{P}2$). In particular, P_1 is a prefix of $\varphi^k(a_t)$, for some $t \leq j$ and some k satisfying

$$|k| \;\leq\; A_j\left(\left|u_0^{-1}\varphi^r(u_0)\right|_H + |r|\right). \tag{18}$$

Let $\pi_0 = u_0 a_i^{-1}$. Since $\varphi^r(\pi_0)$ is the free reduction of $\varphi^r(u_0)\varphi^r(a_i^{-1})$, the common prefix P of π_0 and $\varphi^r(\pi_0)$ is a concatenation of a common prefix of u_0 and $\varphi^r(u_0)$ (hence a prefix of P_0), and a (possibly empty) subword P_i of $\varphi^r(a_i)^{-1}$. We may therefore write $P = P_1 P_3 \cdots P'_{j'} P_i$ for some $j' \leq j$, where $P'_{j'}$ is a prefix of $P_{j'}$. This fits with the structure in ($\mathcal{P}2$), although we still need to determine the bound on k, which we will do below, while we verify ($\mathcal{P}1''$) holds.

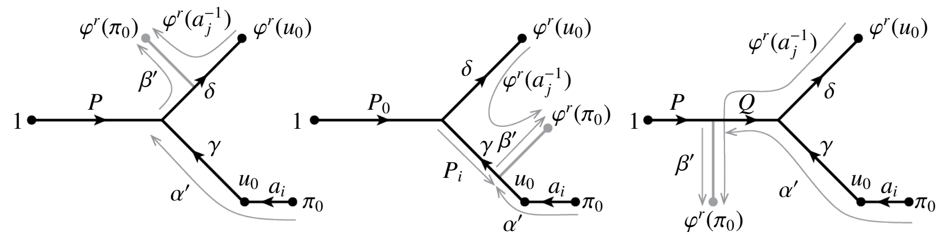

FIGURE 4. Three cases of cancellation between $\varphi^r(a_i^{-1})$ and other words

We seek α' and β' satisfying ($\mathcal{P}1''$), so we need to understand the free reduction of $\pi_0^{-1}\varphi^r(\pi_0)$. We have

$$\pi_0^{-1}\varphi^r(\pi_0) \;=\; a_i u_0^{-1}\varphi^r(u_0)\varphi^r(a_i)^{-1} \;=\; a_i\gamma\delta\varphi^r(a_i)^{-1}.$$

Cancellation in $a_i\gamma\delta\varphi^r(a_i)^{-1}$ can only occur where δ abuts $\varphi^r(a_i)^{-1}$. One of three things can occur. Either

($C1$) $\varphi^r(a_i)^{-1}$ does not cancel into either γ or P_0 (the left diagram of Figure 4),
($C2$) $\varphi^r(a_i)^{-1}$ completely cancels with δ and continues cancelling into γ (the middle diagram in Figure 4), or
($C3$) $\varphi^r(a_i)^{-1}$ completely cancels with δ and continues cancelling into P_0 (the right diagram in Figure 4).

In ($C1$) we can take $\alpha' = a_i\gamma$ and β' to be the free reduction of $\delta\varphi^r(a_i)^{-1}$. Then it is straightforward to see that ($\mathcal{P}1''$) holds.

For ($C2$), we take P_i as above: it is the subword of $\varphi^r(a_i)^{-1}$ that cancels into γ. Then, setting α' to be the free reduction of $a_i\gamma P_i$ and β' to be the free reduction of $P_i^{-1}\delta\varphi^r(a_i)^{-1}$, we can check that ($\mathcal{P}1''$) is satisfied. Take $P = P_0 P_i$.

Finally, for ($C3$), we let Q be the suffix of P_0 that cancels into $\varphi^r(a_i)^{-1}$, and P be the prefix of P_0 so that $P_0 = PQ$ as words. (Note in this case P_i is empty.) Then $\alpha' = a_i\gamma Q^{-1}$ and β' equal to the free reduction of $Q\delta\varphi^r(a_i)^{-1}$ satisfy ($\mathcal{P}1''$).

In each case, the free reduction of $u_0^{-1}\varphi^r(u_0)$ is a product of a subword of the free reduction of $\pi_0^{-1}\varphi^r(\pi_0)$ with a subword of $\varphi^r(a_i)^{\pm 1}$. By ($\mathcal{P}1'''$), $\pi_0^{-1}\varphi^r(\pi_0)$ is equal to $\pi^{-1}\varphi^r(\pi)$, which is a subword of $\alpha\beta$. Hence by Proposition 4.5, $\left|u_0^{-1}\varphi^r(u_0)\right|_H \leq (2m+1)\left(|\alpha\beta|_H + 2|r| + 1\right)$. In each case P has the form required for ($\mathcal{P}2$), and this bound, together with (18) and increasing the value as A_i if necessary, gives the required bound on $|k|$. ◻

7. The inductive structure of $\mathcal{H}$-twisted conjugacy in F

We explained in Section 2 that the conjugacy problem in H amounts to a twisted conjugacy problem in F which can take one of three forms. The most involved of the three is what we refer to as the $\mathcal{H}$-twisted conjugacy problem. We will show here that when this problem has a solution, it has a solution of one of a number of particular forms. The large number of possibilities for this form leads to the following proposition having a somewhat involved statement. But all the subwords are described in terms of the 'constants' p, u_0, v_0, u_1, v_1 or inductively in terms of a word $\hat{x}$ which is a solution to a lower rank instance of the same problem. This will allow us to estimate the lengths of solutions. Those estimates will feed into upper bounds on the conjugator length of H. Also this proposition will mean that solutions to the $\mathcal{H}$-twisted conjugacy problem can found by searching though a polynomially sized family of possibilities. This will feed into polynomial time solutions to the conjugacy and conjugacy-search problems in H.

Proposition 7.1. *Suppose $i \in \{1, \ldots, m\}$, $p > 0$ are integers and u_0, v_0, u_1, v_1, x are reduced words on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$. Suppose x is non-empty word and has rank i. Suppose the concatenations $u_0 x v_0^{-1}$ and $u_1^{-1} x v_1$ are reduced words and satisfy*

$$\varphi^{-p}(u_0 x v_0^{-1}) \ = \ u_1^{-1} x v_1 \ \text{ in } \ F. \tag{19}$$

Then there exists a word X on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$ that satisfies

$$\varphi^{-p}(u_0 X v_0^{-1}) \ = \ u_1^{-1} X v_1 \ \text{ in } \ F \tag{20}$$

and takes the following form. If $i = 1$, then

($\mathcal{X}1$) $X = U_1 U_2 U_3 V$ as words

for some subwords U_1, U_2, and U_3 of $(u_1 \varphi^{-p}(u_0))^{\pm 1}$ and some suffix V of $v_1 \varphi^{-p}(v_0)$.

There exists a constant $C > 0$ such that if $i > 1$, then either

($\mathcal{X}2$) $X = x$ is a subword L of $\varphi^{-p}(u_0)$ or R of $\varphi^{-p}(v_0^{-1})$,
($\mathcal{X}3$) $X = L\,S\ M\,P\,R$ as words, or
($\mathcal{X}4$) $X = x = L\,\hat{x}\,R$ as words,

where

- *L is a subword of $\varphi^{-p}(u_0)$,*
- *S is either*
  - *$S = S_i \cdots S_3\, S_1$ where*
    - *S_1^{-1} is a prefix of $\varphi^k(a_t)$ for some $t \le i$ and $|k| \le C\,(|u_0|_F + |u_1|_F + p)$,*
    - *S_j is a subword of $\varphi^{-p}(a_j)$ for $j = 3, \ldots, i$,*
  - *a subword of $\hat{S}$ where $\varphi^{-p}(\hat{S})$ is a subword of $(u_1 \varphi^{-p}(u_0))^{-1}$, or*
  - *a subword of $\varphi^{-p}(\hat{S})$ where $\hat{S}$ is a subword of $\varphi^{-p}(u_0)$,*
- *$M = M_1 M_2$ or $M_2^{-1} M_1^{-1}$, where*
  - *$M_1 = \pi \varphi^p(\pi) \cdots \varphi^{p(q-1)}(\pi)$, where*
    - *$qp \le C(|u_0|_H + |u_1|_H + |v_0|_H + |v_1|_H + p)$, and*
    - *$\varphi^{-p}(\pi)$ is a concatenation of a subword of $(u_1 \varphi^{-p}(u_0))^{-1}$ with S, or of P with a subword of $(v_1 \varphi^{-p}(v_0))^{-1}$,*
  - *$\varphi^p(M_2)$ is a subword of $\varphi^{-p}(u_0)$ or $\varphi^{-p}(v_0^{-1})$,*
- *P is either*
  - *$P = P_1\, P_3 \cdots P_i$, where*
    - *P_1 is a prefix of $\varphi^{k'}(a_t)$ for some $t \le i$ and $|k'| \le C\,(|v_0|_F + |v_1|_F + p)$,*
    - *P_j is a subword of $\varphi^{-p}(a_j^{-1})$ for $j = 3, \ldots, i$, or*
  - *a subword of $\varphi^{-p}(\hat{P})$ where $\hat{P}$ is a subword of $\varphi^{-p}(v_0^{-1})$,*

  - *a subword of $\hat{P}$ where $\varphi^{-p}(\hat{P})$ is a subword of $v_1\varphi^{-p}(v_0)$,*
- *R is a subword of $\varphi^{-p}(v_0^{-1})$,*
- *$\hat{x}$ has rank $j < i$ and satisfies $\varphi^{-p}(\hat{u}_0\hat{x}\hat{v}_0^{-1}) = \hat{u}_1^{-1}\hat{x}\hat{v}_1$ where*
  - *$\hat{u}_0\hat{u}_1$ is reduced and is a subword of $\varphi^{-p}(u_0)$,*
  - *$\hat{v}_0\hat{v}_1$ is reduced and is a subword of $\varphi^{-p}(v_0)$.*

A curious feature of Proposition 7.1 is that if $u, v \in H$ satisfy $uw = wv$ in H for some $w \in F$ (not just in H), and $w = u_0xv_0^{-1}$ as per the proposition, then $u_0Xv_0^{-1}$, which will also be in F, is another conjugator and also has this 'nice' structure.

In general, this structure leads to a quadratic upper bound on the length (see Lemma 8.1 below). To improve it to a linear upper bound we need to replace the word M_1 appearing in ($\mathcal{X}$3) with s^{qp} (see Lemma 8.4 below). We therefore swap X for a word $\hat{X}$, which unlike X may represent an element of $H \smallsetminus F$. Equations (19) and (20) may therefore not make sense for $\hat{X}$. For the iteration through case ($\mathcal{X}$4), then, we will instead use:

Lemma 7.2. *With the notation from Proposition 7.1, if X has form ($\mathcal{X}$4), and if $\hat{X} \in H$ satisfies*

$$s^p\ \hat{u}_0\ \hat{X}\ \hat{v}_0^{-1}\ s^{-p} \ = \ \hat{u}_1^{-1}\ \hat{X}\ \hat{v}_1 \quad \text{in } H$$

then

$$s^p\ u_0\ L\ \hat{X}\ R\ v_0^{-1}\ s^{-p} \ = \ u_1^{-1}\ L\ \hat{X}\ R\ v_1 \quad \text{in } H. \tag{21}$$

If $\tilde{w} = u_0\ L\ \hat{X}\ R\ v_0^{-1}$, then (21) amounts to $uw = wv$ in H, where $w = \tilde{w}s^r$, $u = \tilde{u}s^p$, $v = \tilde{v}s^p$, $\tilde{u} = u_0u_1$, and $\varphi^{-r}(\tilde{v}) = v_0v_1$.

We will prove this lemma after proving Proposition 7.1. Before we prove either, here is a lemma which is straight-forward, but which we highlight as we will call on it to remove a subword from x.

Lemma 7.3. *Suppose $x = x_0x_1$, $\varphi^{-p}(u_0x_0) = u_1^{-1}x_0$, $\varphi^{-p}(x_1v_0^{-1}) = x_1v_1$, and $\varphi^{-p}(y) = y$ in F. Then $\varphi^{-p}(u_0x_0yx_1v_0^{-1}) = u_1^{-1}x_0yx_1v_1$ in F.*

Proof of Proposition 7.1. We begin with the case $i = 1$. As $\mathrm{rank}(x) = 1$, $\varphi(x) = x$, and so (19) rearranges to the conjugacy relation

$$u_1\varphi^{-p}(u_0)x \ = \ xv_1\varphi^{-p}(v_0) \ \text{ in } \ F. \tag{22}$$

Therefore, by Lemma 2.1, there is *some* $x_0 \in F$ which satisfies (22) in place of x (but may fail to satisfy (19) since φ need not fix x_0) and is the concatenation of some prefix of $(u_1\varphi^{-p}(u_0))^{-1}$ with some suffix of $v_1\varphi^{-p}(v_0)$.

If $u_1\varphi^{-p}(u_0) = 1$, then (20) holds with X the empty word. Assume, then, that $u_1\varphi^{-p}(u_0) \neq 1$. Since both x and x_0 conjugate $u_1\varphi^{-p}(u_0)$ to $v_1\varphi^{-p}(v_0)$ in F, we have $x = \sigma^l x_0$ in F for some integer l and some reduced word σ some power of which freely equals $u_1\varphi^{-p}(u_0)$. If $\mathrm{rank}(\sigma) = 1$, then $\mathrm{rank}(x_0) = 1$ also and so $\varphi(x_0) = x_0$ and (20) holds with $X = x_0$. If, on the other hand, the $\mathrm{rank}(\sigma) \geq 2$, then take $X = x$. Then (20) is (19) and so holds. And, as $\mathrm{rank}(x) = 1$, all of σ^l apart from some prefix σ' of σ (if $l > 0$) or σ^{-1} (if $l < 0$) must cancel into x_0 in $\sigma^l x_0$, and so X is the concatenation of a prefix σ' of σ or σ^{-1} with a suffix of x_0.

Let σ_0 be the maximal suffix of σ such that σ_0^{-1} is prefix of σ. Then there is a subword σ_1 of σ such that for all $c \in \mathbb{Z}$, as words $\sigma^c = \sigma_0^{-1}\sigma_1^c\sigma_0$. So, as some power of σ freely equals $u_1\varphi^{-p}(u_0)$, we have that $\sigma_0^{-1}\sigma_1$ freely equals a prefix of $(u_1\varphi^{-p}(u_0))^{\pm 1}$ and σ_0 freely equals a suffix. So σ freely equals the concatenation of two subwords of $(u_1\varphi^{-p}(u_0))^{\pm 1}$, and the same is true of σ'.

So in each case X is the concatenation of three subwords of $(u_1\varphi^{-p}(u_0))^{\pm 1}$ with a suffix of $v_1\varphi^{-p}(v_0)$, completing the proof in the case $i = 1$.

Now assume $2 \leq i \leq m$. By (19), $\varphi^{-p}(u_0)\varphi^{-p}(x)\varphi^{-p}(v_0^{-1}) = u_1^{-1}xv_1$ in F. The right-hand side is reduced, but there may be cancellation on the left at the start and end of $\varphi^{-p}(x)$. So, after free reduction on the left, a (perhaps empty) subword of $\varphi^{-p}(x)$ remains. This subword will also be a subword of $u_1^{-1}xv_1$. Define Π to be its overlap with x.

If Π is the empty word, then x is a subword of either $\varphi^{-p}(u_0)$ or $\varphi^{-p}(v_1^{-1})$, and so $X = x$ has the form of ($\mathcal{X}2$). So for the remainder of the proof we assume that Π is nonempty.

As Π is a subword of both x and $\varphi^{-p}(x)$, we can define L, R, Y and Z so that

$$x = L\,\Pi\,R \qquad \text{and} \qquad \varphi^{-p}(x) = Y\,\Pi\,Z \qquad \text{as words.} \tag{23}$$

We have

$$u_1^{-1}\,L\,\Pi\,R\,v_1 \;=\; \varphi^{-p}(u_0)\,Y\,\Pi\,Z\,\varphi^{-p}(v_0^{-1}) \qquad \text{in } F$$

by (19). We claim that

($\mathcal{E}1$) L is a subword of $\varphi^{-p}(u_0)$,
($\mathcal{E}2$) R^{-1} is a subword of $\varphi^{-p}(v_0)$,
($\mathcal{E}3$) Y^{-1} is a suffix of the freely reduced form of $u_1\varphi^{-p}(u_0)$, and
($\mathcal{E}4$) Z is a suffix of the freely reduced form of $v_1\varphi^{-p}(v_0)$.

By definition of Π,

$$\varphi^{-p}(u_0)Y \text{ freely reduces to } u_1^{-1}\,L. \tag{24}$$

Also, $\varphi^{-p}(u_0)$ and Y are freely reduced words, so if Y fully cancels into $\varphi^{-p}(u_0)$ on free reduction of $\varphi^{-p}(u_0)Y$, then Y^{-1} is a suffix of $\varphi^{-p}(u_0)$ and (24) gives us ($\mathcal{E}1$) and ($\mathcal{E}3$). If, on the other hand, a non-empty suffix of Y survives free reduction of $\varphi^{-p}(u_0)\,Y$, then L is the empty word: the last letter of L, were there one, would have to have been part of Π since it would also have to have been the last letter of Y. So ($\mathcal{E}1$) trivially holds and ($\mathcal{E}3$) again follows from (24).

To complete the proof, when Π is nonempty we need to explain how to replace Π with some $\hat{\Pi}$ so that $X = L\hat{\Pi}R$ is of the required form.

Let $\pi_1 \cdots \pi_k$ be the rank-i piece decomposition of x. By Lemma 3.2, $\varphi^{-p}(\pi_1)\cdots\varphi^{-p}(\pi_k)$ is the rank-i decomposition of $\varphi^{-p}(x)$ into pieces and, in particular, is reduced.

Since Π is a subword of x, there are integers a, b such that x is a subword of $\pi_a \cdots \pi_b$. We choose a and b so that either

I. $a < b$ and $\Pi = \pi'_a\pi_{a+1}\cdots\pi_{b-1}\pi'_b$, for some (perhaps empty) suffix π'_a of π_a and (perhaps empty) prefix π'_b of π_b, where $\pi'_a \neq \pi_a$ if $1 < a$ and $\pi'_b \neq \pi_b$ if $b < k$, or
II. $a = b$ and $\Pi = \pi'_a$ for a nonempty subword π'_a of π_a.

As Π is a subword of $\varphi^{-p}(x)$, its rank–i pieces line up with those in $\varphi^{-p}(\pi_1)\cdots\varphi^{-p}(\pi_k)$. This tells us

($\mathcal{S}1$) Π is a subword of $\varphi^{-p}(\pi_{a+e})\cdots\varphi^{-p}(\pi_{b+e})$ for some $e \in \mathbb{Z}$.

The value of e will be important.

From (23) and ($\mathcal{E}1$)–($\mathcal{E}4$) we make the following observations.

($\mathcal{S}2$) If $c < a$, then π_c is a subword of $\varphi^{-p}(u_0)$.
($\mathcal{S}3$) If $c > b$, then π_c^{-1} is a subword of $\varphi^{-p}(v_0)$.
($\mathcal{S}4$) If $c < a+e$, then $\varphi^{-p}(\pi_c)^{-1}$ is a subword of the freely reduced form of $u_1\varphi^{-p}(u_0)$.
($\mathcal{S}5$) If $c > b+e$, then $\varphi^{-p}(\pi_c)$ is a subword of the freely reduced form $v_1\varphi^{-p}(v_0)$.

Case I. $\Pi = \pi'_a\pi_{a+1}\cdots\pi_{b-1}\pi'_b$.

We make the following further observations that apply in this case. They follow by considering Π simultaneously as a subword of x, which has piece decomposition $\pi_1\cdots\pi_k$, and of $\varphi^{-p}(x)$, with piece decomposition $\varphi^{-p}(\pi_1)\cdots\varphi^{-p}(\pi_k)$. As mentioned, the breaks between pieces in Π must line up in the two words x and $\varphi^{-p}(x)$.

($\mathcal{S}6$) π'_a is a suffix of $\varphi^{-p}(\pi_{a+e})$ (as well as of π_a),
($\mathcal{S}7$) $\pi_i = \varphi^{-p}(\pi_{i+e})$ for $a+1 \le i \le b-1$,
($\mathcal{S}8$) π'_b is a prefix of $\varphi^{-p}(\pi_{b+e})$ (as well as of π_b).

We will show that the proposition is satisfied with X as per ($\mathcal{X}3$). We will divide into three subcases according to the value of e. Having e positive is similar to e negative. Indeed, taking inverses of both sides of equation (19) interchanges the roles of u_0 and u_1 with v_0 and v_1, respectively, and puts x^{-1} in place of x. If $\hat{\pi}_i$ is the i–th piece in the rank i piece decomposition of x^{-1}, then $\hat{\pi}_i = \pi_{k-i}^{-1}$, and this means that to satisfy ($\mathcal{S}7$) with $\hat{\pi}_i$ instead, we use $-e$ instead of e. So in Cases Ib and Ic below, when e is negative, swapping the roles of u_i and v_i accordingly will give the structure of X^{-1}, and that of X will then be apparent.

Case Ia. When $e = 0$.

If we remove the pieces $\pi_{a+1},\ldots,\pi_{b-1}$ from x leaving $X_0 := L\pi'_a\pi'_bR$, then $\varphi^{-p}(u_0X_0v_0^{-1}) = u_1^{-1}X_0v_1$ in F by Lemma 7.3. Next we will replace π'_a with $\hat{\pi}'_a$, and π'_b with $\hat{\pi}_b$, which come from Corollary 6.2, as explained below, giving $X := L\hat{\pi}'_a\hat{\pi}'_bR$. We will show this too satisfies (20).

As π'_a is a common suffix of π_a and $\varphi^p(\pi_a)$ (and so a common prefix of π_a^{-1} and $\varphi^{-p}(\pi_a)^{-1}$), Corollary 6.2 gives us a word $\hat{\pi}_a$ so that

($\mathcal{A}1$) if $\alpha\beta$ is the free reduction of $\pi_a\varphi^{-p}(\pi_a)^{-1}$, with $\pi_a = \alpha\pi'_a$ and $\varphi^{-p}(\pi_a) = \beta^{-1}\pi'_a$ as words, then $\alpha\beta$ is also the free reduction of $\hat{\pi}_a\varphi^{-p}(\hat{\pi}_a)^{-1}$ in F, with $\hat{\pi}_a = \alpha\hat{\pi}'_a$ and $\varphi^{-p}(\hat{\pi}_a) = \beta^{-1}\hat{\pi}'_a$ as words, and
($\mathcal{A}2$) the common suffix $\hat{\pi}'_a$ of $\hat{\pi}_a$ and $\varphi^{-p}(\hat{\pi}_a)$ is of the form $\hat{\pi}'_a = S_i\cdots S_3S_1$, where
  - S_1^{-1} is a prefix of $\varphi^k(a_t)$ for some $t \le i$ and $|k| \le A_i\left(\left|\pi_a\varphi^{-p}(\pi_a)^{-1}\right|_H + p\right)$,
  - S_j is a subword of $\varphi^{-p}(a_j)$ for $j = 3,\ldots,i$.

First we check that $S = S_i\cdots S_3S_1$ fits the scheme of the proposition. To do this we need to bound $\left|\pi_a\varphi^{-p}(\pi_a)^{-1}\right|_H$ so that ($\mathcal{A}2$) leads to the required bound on $|k|$. We could use ($\mathcal{E}1$) and ($\mathcal{E}3$) alongside Proposition 4.5, but we can do better as follows. By (24), $LY^{-1} = u_1\varphi^{-p}(u_0)$ in F. The last letters of L and Y must be different, since otherwise that letter could be added into Π. So LY^{-1} is reduced. The free reduction of $\pi_a(\pi'_a)^{-1}$ is a suffix of L, and that of $\varphi^{-p}(\pi_a)(\pi'_a)^{-1}$ is a suffix of Y. Hence the free reduction of $\pi_a\varphi^{-p}(\pi_a)^{-1}$ is a subword of the free reduction of $u_1\varphi^{-p}(u_0)$. So Proposition 4.5 gives the first inequality of:

$$\left|\pi_a\varphi^{-p}(\pi_a)^{-1}\right|_H \;\leq\; (2m+1)\left|u_1\varphi^{-p}(u_0)\right|_H \;\leq\; (2m+1)\left(|u_0|_F+|u_1|_F+2p\right).$$

We deduce that there is a constant $C>0$ such that

$$|k| \;\leq\; C\left(|u_0|_F+|u_1|_F+p\right).$$

Similarly, π'_b is a common prefix of π_b and $\varphi^{-p}(\pi_b)$, so Corollary 6.2, tells us that there is word $\hat{\pi}_b$ such that

($\mathcal{B}1$) if $\gamma\delta$ is the free reduction of $\pi_b^{-1}\varphi^{-p}(\pi_b)$, with $\pi_b = \pi'_b\gamma^{-1}$ and $\varphi^{-p}(\pi_b) = \pi'_b\delta$ as words, then $\gamma\delta$ is also the free reduction of $\hat{\pi}_b^{-1}\varphi^{-p}(\hat{\pi}_b)$ in F, with $\hat{\pi}_b = \hat{\pi}'_b\gamma^{-1}$ and $\varphi^{-p}(\hat{\pi}_b) = \hat{\pi}'_b\delta$ as words, and

($\mathcal{B}2$) the common prefix $\hat{\pi}'_b$ of $\hat{\pi}_b$ and $\varphi^{-p}(\hat{\pi}_b)$ is of the form $\pi'_b = P_1P_3\cdots P_i$, where
- P_1 is a prefix of $\varphi^{k'}(a_t)$ for some $t\leq i$ and $|k'|\leq A_i\left(\left|\pi_b^{-1}\varphi^{-p}(\pi_b)\right|_H+p\right)$,
- P_j is a subword of $\varphi^{-p}(a_j^{-1})$ for $j=3,\ldots,i$.

Similar reasoning to the above give us $|k'|\leq C(|v_0|_F+|v_1|_F+p)$. It follows that $X = LSPR$ has form ($\mathcal{X}3$), with M the empty word.

To conclude, we need to prove X will satisfy (20). First we work on the left. Using the notation from ($\mathcal{A}1$), we have $L = \pi_1\cdots\pi_{a-1}\alpha$ and $Y = \varphi^{-1}(\pi_1\cdots\pi_{a-1})\beta^{-1}$, so $Y = \varphi^{-p}(L\alpha^{-1})\beta^{-1}$. Then by (24), $\varphi^{-p}(u_0L\alpha^{-1})\beta^{-1} = u_1^{-1}L$. Rearranging and using that $\alpha\hat{\pi}'_a = \hat{\pi}_a$ and $\beta^{-1}\hat{\pi}_a = \varphi^{-p}(\pi_a)$, we get

$$\varphi^{-p}(u_0L\hat{\pi}'_a) \;=\; u_1^{-1}L\beta\varphi^{-p}(\alpha\hat{\pi}'_a) \;=\; u_1^{-1}L\beta\varphi^{-p}(\hat{\pi}_a) \;=\; u_1^{-1}L\beta\beta^{-1}\hat{\pi}'_a \;=\; u_1^{-1}L\hat{\pi}'_a.$$

Similar calculations on the right yield $\varphi^{-p}(\hat{\pi}'_bRu_0^{-1}) = \hat{\pi}'_bRv_1$. The left and right, working together, give us (20).

Case Ib. When $|e| > b-a-1$.

First assume $e>0$. Write $\Pi = SMP$, where $S = \pi'_a$, $M = M_2 = \pi_{a+1}\cdots\pi_{b-1}$, and $P = \pi'_b$. Then $M_2 = \varphi^{-p}(\pi_{a+1+e}\cdots\pi_{b-1+e})$ by ($\mathcal{S}7$), and $\varphi^p(M_2) = \pi_{a+1+e}\cdots\pi_{b-1+e}$ is a subword of $\varphi^{-p}(v_0)^{-1}$ by ($\mathcal{E}2$). Meanwhile, S is a subword of π_a, and $\varphi^{-p}(\pi_a)$ is a subword of $(u_1\varphi^{-p}(u_0))^{-1}$ by ($\mathcal{S}4$). Finally, P is a subword of $\varphi^{-p}(\pi_{b+e})$, and $\hat{P} = \pi_{b+e}$ is a subword of $\varphi^{-p}(v_0^{-1})$ by ($\mathcal{S}3$).

Now assume that $e<0$. Then the above gives the structure for Π^{-1}, after swapping the roles of u_i and v_i. Hence we get $\Pi = SMP$, where S is a subword of $\varphi^{-p}(\hat{S})$ where $\hat{S}$ is a subword of $\varphi^{-p}(u_0)$, $M = M_2$ and $\varphi^{-p}(M_2)$ is a subword of $v_1\varphi^{-p}(v_0)$, and P is a subword of $\hat{P}$, and $\varphi^{-p}(\hat{P})$ is a subword of $v_1\varphi^{-p}(v_0)$.

Case Ic. When $0<|e|\leq b-a-1$.

Suppose $e>0$. We have

$$\Pi \;=\; \pi'_a\pi_{a+1}\cdots\pi_{b-1}\pi'_b.$$

The number of pieces in Π aside from π'_a and π'_b is $b-a-1$. Let q be the maximal integer such that $qe\leq b-a-1$. Let

$$\pi \;=\; \pi_{a+1}\cdots\pi_{a+e}.$$

By repeated applications of ($\mathcal{S}7$),

$$\Pi = SMP \quad \text{in } F$$

where $M = M_1M_2$ and

$$\begin{aligned} S &= \pi'_a \\ M_1 &= \pi\varphi^p(\pi)\cdots\varphi^{p(q-1)}(\pi) \\ M_2 &= \varphi^{pq}(\pi_{a+1}\cdots\pi_{b-1-qe}) \\ P &= \pi'_b. \end{aligned}$$

Then S and P are as in Case Ib. Notice that $\varphi^{-p}(\pi)$ is a concatenation of a subword of Y (so, by ($\mathcal{E}3$), a subword of the freely reduced form of $(u_1\varphi^{-p}(u_0))^{-1}$) with $S = \pi'_a$. By ($\mathcal{S}7$), $\varphi^p(M_2) = \pi_{a+1+(q+1)e}\cdots\pi_{b-1+e}$. Note that $a+1+(q+1)e \neq b$, by our choice of q. So $a+1+(q+e) > b$, which implies $\varphi^p(M_2)$ is a subword of $\varphi^{-p}(v_0^{-1})$ by ($\mathcal{E}2$).

We now begin the work necessary to establish the required bound on q.

First, for $a \leq c \leq c' < a+e$, we have that $\varphi^{-p}(\pi_c\cdots\pi_{c'})$ is a subword of Y, and so of the freely reduced form of $(u_1\varphi^{-p}(u_0))^{-1}$ by ($\mathcal{E}3$). Also $\varphi^{-p}(\pi_{a+e})$ is a product of subwords of $(u_1\varphi^{-p}(u_0))^{-1}$ and π_a. So, by Proposition 4.5, there is a constant $C > 0$, depending only on m, such that

$$|\pi_c\cdots\pi_{c'}|_H \;\leq\; C\left(\left|u_1\varphi^{-p}(u_0)\right|_H + p\right) \quad \text{for } a \leq c \leq c' \leq a+e. \tag{25}$$

Similarly,

$$|\pi_d\cdots\pi_{d'}|_H \;\leq\; C\left(\left|v_1\varphi^{-p}(v_0)\right|_H + p\right) \quad \text{for } b \leq d \leq d' \leq b+e. \tag{26}$$

We now bound qp. Let $a < c \leq a+e$ and choose $b \leq d < b+e$ so that $\pi_d = \varphi^{q'p}(\pi_c)$, where q' is q or $q+1$ according to whether $c > b-1-qe$ or not. Suppose $\varphi(\pi_c) \neq \pi_c$ and the rank of π_c is $j \leq i$. We will establish an upper bound on $q'p$ by combining an upper bound on $|\pi_d|_H$ from (26) with an understanding of how fast π_c can grow under iterates of φ^{-1} from Proposition 4.4.

We claim that we can choose such c so that $j = i$ (that is, so that π_c has rank i). First assume that $e = 1$. Then ($\mathcal{S}7$) gives $\pi_c = \varphi^{-p}(\pi_{c+1})$. In particular, both π_c and π_{c+1} have the same rank, which must therefore be i, since in a pair of adjacent pieces of a rank–i decomposition of a word, at least one must have rank i. So we may assume $e > 1$. Since we assume $\varphi(\pi_c) \neq \pi_c$, its rank satisfies $j \geq 2$ by Lemma 4.1. Suppose $j < i$. Then both the the neighbours π_{c+1} and π_{c-1} of π_c must have rank i. But $i > j \geq 2$, so $\varphi(\pi_{c+1}) \neq \pi_{c+1}$ and $\varphi(\pi_{c-1}) \neq \pi_{c-1}$ by Lemma 4.1 again. The inequality $a < c \pm 1 \leq a+e$ holds for at least one of $c+1$ or $c-1$. We replace c with the corresponding number and may therefore assume $j = i$.

If $(q'p)^{i-1} < \frac{1}{C_i}|\pi_c|_F$, then $C_i^{\frac{1}{i-1}}q'p \;<\; |\pi_c|_F^{\frac{1}{i-1}}$, and so Corollary 4.9 and then (25) imply

$$C_i^{\frac{1}{i-1}}q'p \;<\; K_i^{\frac{1}{i-1}}\,|\pi_c|_H \;\leq\; K_i^{\frac{1}{i-1}}C\left(\left|u_1\varphi^{-p}(u_0)\right|_H + p\right).$$

If, on the other hand, $(q'p)^{i-1} \;\geq\; \frac{1}{C_i}|\pi_c|_F$, then, we may apply Proposition 4.4 to $\pi_d = \varphi^{q'p}(\pi_c)$, giving

$$|\pi_d|_F \;=\; \left|\varphi^{q'p}(\pi_c)\right|_F \;\geq\; \left(C_i^{\frac{1}{i-1}}q'p - |\pi_c|_F{}^{\frac{1}{i-1}}\right)^{i-1}.$$

Rearranging and then combining this with Corollary 4.9 and inequalities (25) and (26) yields

$$C_i^{\frac{1}{i-1}} q' p \;\leq\; K_i^{\frac{1}{i-1}} \left(|\pi_c|_H + |\pi_d|_H\right) \;\leq\; K_i^{\frac{1}{i-1}} C \left(\left|u_1 \varphi^{-p}(u_0)\right|_H + \left|v_1 \varphi^{-p}(v_0)\right|_H + 2p \right).$$

Increasing the value of C if necessary, we then get

$$qp \;\leq\; C(|u_0|_H + |u_1|_H + |v_0|_H + |v_1|_H + p). \tag{27}$$

So, provided there is some such c with $\varphi(\pi_c) \neq \pi_c$, we have the required bound on qp, and taking $\hat{\Pi} = \Pi = SMP$ gives an X satisfying ($\mathcal{X}3$).

If, on the other hand, π_c is fixed by φ for all $a < c \leq a+e$ then we may cut a big chunk out of Π, since then $\varphi^p(\pi) = \pi$ and $M = \pi^{pq} \varphi^{pq}(\pi_a \cdots \pi_{b-1-qe})$. By Lemma 7.3 we can remove π^{pq}, and then $\hat{\Pi} = SMP$, where $M = M_2$, gives an X satisfying the proposition.

When $e < 0$, we get the structure of Π^{-1} and of $\hat{\Pi}^{-1}$ from the above argument, once the roles of u_i and v_i have been swapped. As above, S and P will be as in Case Ib. For M, we need to change the order of M_1 and M_2, but we have to be careful as taking the inverse of M_1 changes its structure. The easiest way to express this is to say $M = M_2^{-1} M_1^{-1}$, with M_1 and M_2 obtained as above, but with the u_i and v_i exchanged.

Case II. $\Pi = \pi'_a$.

By ($\mathcal{S}1$), Π is a subword of $\varphi^{-p}(\pi_{a+e})$. If $e < 0$, then by ($\mathcal{S}2$), π_{a+e} is a subword of $\varphi^{-p}(u_0)$. So $\hat{\Pi} = \Pi$ satisfies the conditions of form ($\mathcal{X}3$) of the proposition with $S = \Pi$, $\hat{S} = \pi_{a+e}$, and M and P both the empty word. If $e > 0$, then we take $\hat{\Pi} = \Pi$ and it similarly satisfies the conditions with S and M both the empty word and $P = \Pi$, which is a subword of $\varphi^{-p}(\pi_{a+e})$ and $\hat{P} = \pi_{a+e}$ is a subword of $\varphi^{-p}(v_0^{-1})$.

On the other hand, assume $e = 0$. If Π is either a prefix or a suffix of π_a then we can apply Corollary 6.2. If Π is a prefix, we replace it with $\hat{\Pi} = P = P_1 P_3 \cdots P_i$, as in Corollary 6.2. The proof that (20) holds, and of the bound on the k in P_1 are the same as for Case Ia (treating π'_a as the empty word and $b = a+1$). If Π is instead a suffix, replace it with $\hat{\Pi} = S = S_i \cdots S_3 S_1$ using Corollary 6.2, and a similar check gives (20) and a corresponding bound on k'.

What remains is to consider when Π is not a prefix or suffix of π_a. Then it has rank $j < i$. This is where the form ($\mathcal{X}4$) occurs. As reduced words, write $\pi_a = \hat{u}_0 \hat{x} \hat{v}_0^{-1}$ and $\varphi^{-p}(\pi_a) = \hat{u}_1^{-1} \hat{x} \hat{v}_1$, where $\hat{x} = \pi'_a = \Pi$. So $\operatorname{rank}(\hat{x}) < i$ and $\varphi^{-p}(\hat{u}_0 \hat{x} \hat{v}_0^{-1}) = \hat{u}_1^{-1} \hat{x} \hat{v}_1$, as required. Since $\hat{u}_0$ is a suffix of L and $\hat{u}_1$ is a prefix of Y^{-1}, and there is no cancellation between L and Y^{-1} by the definition of Π, we can deduce from (24) that $\hat{u}_0 \hat{u}_1$ is reduced and is a subword of $\varphi^{-p}(u_0)$. □

Proof of Lemma 7.2. By hypothesis

$$s^p \; \hat{u}_0 \; \hat{X} \; \hat{v}_0^{-1} \; s^{-p} \;=\; \hat{u}_1^{-1} \; \hat{X} \; \hat{v}_1 \quad \text{in } H. \tag{28}$$

By (24), $s^p \; u_0 \; s^{-p} \; Y \; \hat{u}_1 = u_1^{-1} L \hat{u}_1$ in H, which together with (28) gives

$$(s^p \; u_0 \; s^{-p} \; Y \; \hat{u}_1)\,(s^p \; \hat{u}_0 \; \hat{X} \; \hat{v}_0^{-1} \; s^{-p}) \;=\; (u_1^{-1} \; L \; \hat{u}_1)\,(\hat{u}_1^{-1} \; \hat{X} \; \hat{v}_1) \quad \text{in } H. \tag{29}$$

By hypothesis, X has form ($\mathcal{X}4$), and as per Case II of our proof of Proposition 7.1, $L = \pi_1 \cdots \pi_{a-1} \hat{u}_0$ and $Y = \varphi^{-p}(\pi_1 \cdots \pi_{a-1}) \hat{u}_1^{-1}$. Comparing these expressions for L and Y we get that in H we have $L = s^{-p} Y \hat{u}_1 s^p \hat{u}_0$, the right-hand side of which is a substring of the left side of (29). So substituting accordingly into the left and cancelling the $\hat{u}_1 \hat{u}_1^{-1}$ from the right, we get

$$s^p \; u_0 \; L \; \hat{X} \; \hat{v}_0^{-1} \; s^{-p} = u_1^{-1} \; L \;\; \hat{X} \; \hat{v}_1 \quad \text{in } H. \tag{30}$$

In the same manner as we derived (30) from (28) using (24), calculations on the right-hand ends using the equation $Z\varphi^{-p}(v_0^{-1}) = Rv_1$ will derive (21) from (30).

The final part of the lemma follows from the discussion in Section 2 or by direct calculation. □

8. Solving the $\mathcal{H}$-twisted conjugacy problem

Lemma 8.1. *There exists a constant $K > 0$, depending only on m, such that the lengths in H of the elements U_1, U_2, U_3, V, L, S, P, M_2, R, S, $\hat{u}_0$, $\hat{u}_1$, $\hat{v}_0$, $\hat{v}_1$, and π that arise in Proposition 7.1 are all at most $K\Sigma$, where*

$$\Sigma \;=\; (|u_0|_H + |u_1|_H + |v_0|_H + |v_1|_H + p).$$

Moreover, $|M_1|_H \leq K\Sigma^2$.

Proof. The $K\Sigma$ upper bounds all follow from applying Proposition 4.5 to the descriptions of the words and the associated bounds given in Proposition 7.1, noting that $\left|\varphi^j(w)\right|_H \leq 2|j| + |w|_H$ and that the i in the proposition is at most m.

For the bound on the length of M_1, observe that

$$\begin{aligned} M_1 \;&=\; \pi\varphi^p(\pi)\cdots\varphi^{p(q-1)}(\pi) \\ &=\; \pi(s^{-p}\pi s^p)(s^{-2p}\pi s^{2p})\cdots(s^{-p(q-1)}\pi s^{p(q-1)}) \\ &=\; (\pi s^{-p})^q s^{qp}. \end{aligned}$$

Combining (27) with $|\pi|_H \leq K\Sigma$ and adjusting K suitably gives $|M_1|_H \leq K\Sigma^2$. □

Recall that the $\mathcal{H}$-twisted conjugacy problem asks: given reduced words $\tilde{u}, \tilde{v}$ on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$ and an integer $p > 0$, do there exist $0 \leq r < p$ and words $x, u_0, v_0, u_1, v_1 \in F$ such that $\tilde{u} = u_0 u_1$ and $\varphi^{-r}(\tilde{v}) = v_0 v_1$, as words, and

$$\varphi^{-p}(u_0 x v_0^{-1}) \;=\; u_1^{-1} x v_1 \;\text{ in }\; F?$$

Lemma 8.2. *For all $i = 1, \ldots, m$, there exists an algorithm that, with input any $(p, \tilde{u}, \tilde{v})$ for which the $\mathcal{H}$-conjugacy problem has a solution $(r, x, u_0, v_0, u_1, v_1)$ with* $\operatorname{rank}(x) = i$*, will exhibit* some *solution $(r, X, u_0, v_0, u_1, v_1)$; the running time of this algorithm is bounded above by a polynomial in $p + |\tilde{u}|_H + |\tilde{v}|_H$ (where the implied constants depend only on the rank m of F).*

Proof. In the following, when we refer to *polynomial bounds*, we will always mean upper bounds that are polynomial in $p + |\tilde{u}|_H + |\tilde{v}|_H$. We induct on i.

Proposition 7.1 tells us that in the case $i = 1$, there is solution $(r, X, u_0, v_0, u_1, v_1)$ in which X takes the form ($\mathcal{X}1$). We can find one such solution in polynomial time as follows. We list all the (r, u_0, u_1, v_0, v_1) such that $0 \leq r < p$ and $\tilde{u} = u_0 u_1$ and $\varphi^{-r}(\tilde{v}) = v_0 v_1$ as words—there are polynomially many and the words involved all have polynomially bounded length. For each we list all U_1, U_2, U_3 and V as per ($\mathcal{X}1$)—again, polynomially many possibilities—and we check whether of not $\varphi^{-p}(u_0 X v_0^{-1}) = u_1^{-1} X v_1$ in F for $X = U_1 U_2 U_3 V$.

Now suppose that there exists a solution in which $\mathrm{rank}(x) = i > 1$ and that the lemma holds when there exists a solutions in which x has lower rank.

Again list the polynomially many (r, u_0, u_1, v_0, v_1) such that $0 \le r < p$ and $\tilde{u} = u_0 u_1$ and $\varphi^{-r}(\tilde{v}) = v_0 v_1$ as words. For each, list the polynomially many words X of the form ($\mathcal{X}2$) or ($\mathcal{X}3$) of Proposition 7.1 and check whether $\varphi^{-p}(u_0 X v_0^{-1}) = u_1^{-1} X v_1$ in F. If this fails to turn up a solution $(r, X, u_0, v_0, u_1, v_1)$, then the proposition tells us that there must be one in which X has the form ($\mathcal{X}4$). Accordingly, for each of the (r, u_0, u_1, v_0, v_1), list all the polynomially many $(L, R, \hat{u}_0, \hat{u}_1, \hat{v}_0, \hat{v}_1)$ satisfying the conditions of Proposition 7.1. (These $L, R, \hat{u}_0, \hat{u}_1, \hat{v}_0, \hat{v}_1$ have polynomially bounded length.)

We are considering polynomially many (r, u_0, u_1, v_0, v_1), and for each one there are polynomially many $(L, R, \hat{u}_0, \hat{u}_1, \hat{v}_0, \hat{v}_1)$, so this amounts to polynomially many possibilities in total. For one of them, there is an $\hat{x}$ with $\mathrm{rank}(\hat{x}) < i$ such that

$$\varphi^{-p}(\hat{u}_0 \hat{x} \hat{v}_0^{-1}) = \hat{u}_1^{-1} \hat{x} \hat{v}_1 \quad \text{in} \quad F$$

and

$$\varphi^{-p}(u_0 L \hat{x} R v_0^{-1}) \;=\; u_1^{-1} L \hat{x} R v_1 \quad \text{in} \quad F.$$

By induction we have a polynomial time algorithm which we can run (in polynomial time overall) on every one of these possibilities, and for one of them it will exhibit some $\hat{X}$ such that

$$\varphi^{-p}(\hat{u}_0 \hat{X} \hat{v}_0^{-1}) = \hat{u}_1^{-1} \hat{X} \hat{v}_1 \quad \text{in} \quad F.$$

So Lemma 7.2 gives us that $\varphi^{-p}(u_0 L \hat{X} R v_0^{-1}) = u_1^{-1} L \hat{X} R v_1$ in F, and thereby we get a solution $(r, X, u_0, v_0, u_1, v_1)$ where $X = L\hat{X}R$. □

Corollary 8.3 ($\mathcal{H}$-twisted conjugacy complexity)**.** *There is an algorithm that takes as input an integer $p > 0$ and reduced words $\tilde{u}$ and $\tilde{v}$ on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$ and determines whether or not there exists a solution $(r, x, u_0, v_0, u_1, v_1)$ to the $\mathcal{H}$-twisted conjugacy problem. If a solution exists, it exhibits one. The running time of the algorithm is bounded above be a polynomial function of $p + |\tilde{u}|_H + |\tilde{v}|_H$.*

Proof. Run the algorithms of Lemma 8.2 for $i = 1, \ldots, m$ on input $(p, \tilde{u}, \tilde{v})$. The time that it takes each to halt is bounded above by a polynomial in $p + |\tilde{u}|_H + |\tilde{v}|_H$. If there exists a solution, one of them will exhibit it. □

Lemma 8.1 gives us many of the ingredients for the desired linear upper bound on the conjugator length of H, but we will need a way around the quadratic bound on $|M_1|_H$. Accordingly, we will manipulate the form of the conjugator in the case ($\mathcal{X}3$) of Proposition 7.1, which is where M_1 appears.

Lemma 8.4. *Suppose $u = \tilde{u}s^p$ and $v = \tilde{v}s^p$ are conjugate elements of H, and there is a solution $(r, x, u_0, v_0, u_1, v_1)$ to the $\mathcal{H}$-twisted conjugacy problem for $(p, \tilde{u}, \tilde{v})$ in which x has form ($\mathcal{X}3$). Let q be as in Proposition 7.1 (in the form of M_1). Then either $u_0 LS\, s^{pq} M_2 P R v_0^{-1} s^r$ or $u_0 LS\, M_2^{-1} s^{-pq} P R v_0^{-1} s^r$ conjugates u to v.*

Proof. We are in the setting of Case Ic of our proof of Proposition 7.1. Assume that $e > 0$. Let $w = u_0 LS\, s^{pq} M_2 P R v_0^{-1} s^r$. We will show that $uw = wv$ in H.

We have that $u = u_0 u_1 s^p$ and $v = \varphi^r(v_0 v_1) s^p$ in H. Also $M_2 = \varphi^{pq}(\pi_{a+1}\cdots\pi_{b-1-qe})$, which equals $\pi_{a+1+qe}\cdots\pi_{b-1}$ by ($\mathcal{S}7$). Plugging this and the other ingredients into w, we get

$$w \;=\; u_0\, \pi_1\cdots\pi_a\; s^{pq}\; \pi_{a+1+qe}\cdots\pi_k\; v_0^{-1}\; s^r.$$

Then, by repeatedly applying ($\mathcal{S}7$) and the identity $s^{-1}gs = \varphi(g)$ for $g \in F$, we get

$$\begin{aligned}
uw &= u_0\; u_1\; s^p\; u_0\; \pi_1\cdots\pi_a\; s^{pq}\; \pi_{a+1+qe}\cdots\pi_k\; v_0^{-1}\; s^r\\
&= u_0\; u_1\; \varphi^{-p}(u_0\;\pi_1\cdots\pi_a)\; s^{p(q+1)}\; \pi_{a+1+qe}\cdots\pi_k\; v_0^{-1}\; s^r\\
&= u_0\; u_1\; \varphi^{-p}(u_0\;\pi_1\cdots\pi_a)\; \varphi^{-p(q+1)}(\pi_{a+1+qe}\cdots\pi_{b-1})\; s^{p(q+1)}\; \pi_b\cdots\pi_k\; v_0^{-1}\; s^r\\
&= u_0\; u_1\; \varphi^{-p}(u_0\;\pi_1\cdots\pi_a)\; \varphi^{-p}(\pi_{a+1}\cdots\pi_{b-1-qe})\; s^{p(q+1)}\; \pi_b\cdots\pi_k\; v_0^{-1}\; s^r\\
&= u_0\; u_1\; \varphi^{-p}(u_0\;\pi_1\cdots\pi_{b-1-qe})\; s^{p(q+1)}\; \pi_b\cdots\pi_k\; v_0^{-1}\; s^r\\
&= u_0\; u_1\; \varphi^{-p}(u_0\;\pi_1\cdots\pi_{b-1-qe})\; \varphi^{-p(q+1)}(\pi_b\cdots\pi_{a+(q+1)e})\; s^{p(q+1)}\; \pi_{a+(q+1)e+1}\cdots\pi_k\; v_0^{-1}\; s^r\\
&= u_0\; u_1\; \varphi^{-p}(u_0\;\pi_1\cdots\pi_{b-1-qe}\;\pi_{b-qe}\cdots\pi_{a+e})\; s^{p(q+1)}\; \pi_{a+(q+1)e+1}\cdots\pi_k\; v_0^{-1}\; s^r.
\end{aligned}$$

By (24), $\varphi^{-p}(u_0\,\pi_1\cdots\pi_{a+e}) = u_1^{-1}\pi_1\cdots\pi_a$. Hence

$$uw = u_0\;\pi_1\cdots\pi_a\; s^{p(q+1)}\; \pi_{a+(q+1)e+1}\cdots\pi_k\; v_0^{-1}\; s^r. \tag{31}$$

Similar calculations give:

$$\begin{aligned}
wv &= u_0\;\pi_1\cdots\pi_a\; s^{pq}\; \pi_{a+1+qe}\cdots\pi_k\; v_0^{-1}\; s^r\; \varphi^r(v_0v_1)\; s^p\\
&= u_0\;\pi_1\cdots\pi_a\; s^{pq}\; \pi_{a+1+qe}\cdots\pi_k\; v_0^{-1}\; v_0\; v_1\; s^{p+r}\\
&= u_0\;\pi_1\cdots\pi_a\; s^{p(q+1)}\; \varphi^p(\pi_{a+1+qe}\cdots\pi_{b-1}\;\pi_b\cdots\pi_k\; v_1)\; s^r.
\end{aligned}$$

The corresponding fact to (24) concerning R and Z is that $Z\varphi^{-p}(v_0^{-1})$ freely reduces to Rv_1. It implies $\varphi^{-p}(\pi_{b+e}\cdots\pi_k\, v_0^{-1}) = \pi_b\cdots\pi_k v_1$. Together with one final application of ($\mathcal{S}7$) to $\varphi^p(\pi_{a+1+qe}\cdots\pi_{b-1})$, this gives

$$wv \;=\; u_0\;\pi_1\cdots\pi_a\; s^{p(q+1)}\; \pi_{a+1+(q+1)e}\cdots\pi_{b+e-1}\;\pi_{b+e}\cdots\pi_k\; v_0^{-1}\; s^r,$$

which equals uw by equation (31).

The proof when $e < 0$ is similar, giving $uw = wv$ in H for $w = u_0 LS\, M_2^{-1} s^{-pq} PRv_0^{-1} s^r$. □

9. Completing our proof of Theorem 1

We will establish a linear upper bound on the conjugator length of H. Suppose u, v and w are words on $a_1^{\pm1}, \ldots, a_m^{\pm1}, s^{\pm1}$ such that $uw = wv$. We will show that there is a word W on $a_1^{\pm1}, \ldots, a_m^{\pm1}, s^{\pm1}$ such that $uW = Wv$ and $\ell(W)$ at most a constant times $|u|_H + |v|_H$.

Write the normal forms of u, v and w as $\tilde{u}s^p$, $\tilde{v}s^p$ and $\tilde{w}s^r$, respectively.

Following the discussion of Section 2, if $p = 0$ then we are in the 0-twisted conjugacy case, and we find W via Proposition 5.1 (I).

When $p \neq 0$, as we can replace u and v by their inverses if necessary, we may assume $p > 0$. As discussed in Section 2 we may also replace w with $u^j w$ so we can assume $0 \le r < p$. By Proposition 4.5, $|u'|_H \le (2m+1)\,|u|_H$ and $|v'|_H \le (2m+1)\,|\varphi^{-r}(\tilde{v})|_H \le (2m+1)(2r+|v|_H)$ for any subwords u' of $\tilde{u}$ and v' of $\varphi^{-r}(\tilde{v})$. We also have $0 \le r < p \le |u|_H$. So it will suffice to bound the length of W in terms of p and of lengths in H of subwords of $\tilde{u}$ and $\varphi^{-r}(\tilde{v})$.

Our u, v and w form either the I- or $\mathcal{H}$-configuration of Figure 1.

In the case of the I-configuration, $w = u_0 v_0^{-1} s^r$ where u_0 and v_0 are prefixes of $\tilde{u}$ and $\varphi^{-r}(\tilde{v})$. Then $|w|_H \leq |u_0|_H + |v_0|_H + r$, and so $W = w$ will be a conjugator which, by the discussion above, satisfies the required length bound.

Now consider the case of the $\mathcal{H}$-configuration. Proposition 7.1 tells us that we have a conjugator $\tilde{w} s^r$, where $\tilde{w} = u_0 X v_0^{-1}$, with u_0 and v_0 prefixes of $\tilde{u}$ and $\varphi^{-v}(\tilde{v})$ respectively, and the form of X following one of ($\mathcal{X}$1)–($\mathcal{X}$4). It suffices for us to show $\ell(X)$ is at most a constant times $|u|_H + |v|_H$. Lemma 8.1 would give this bound but for M_1 in case ($\mathcal{X}$3) and $\hat{x}$ in case ($\mathcal{X}$4).

As remedy, in the event of ($\mathcal{X}$3), we use the conjugator from Lemma 8.4. As L, S, M_2, P, and R are bounded as required, and qp is bounded by (27), the required bound on $|w|_H$ follows.

In the event of case ($\mathcal{X}$4), we iterate this process. By Proposition 7.1 we know that $\mathrm{rank}(\hat{x}) < \mathrm{rank}(x)$ and $\hat{x}$ occurs in a solution to an $\mathcal{H}$-twisted conjugacy problem, namely

$$\varphi^{-p}(\hat{u}_0 \hat{x} \hat{v}_0^{-1}) \ = \ \hat{u}_1^{-1} \hat{x} \hat{v}_1 \tag{32}$$

where $\hat{u}_0$, $\hat{u}_1$, $\hat{v}_0$, and $\hat{v}_1$ are words whose lengths in H are at most a constant multiple of $|u|_H + |v|_H$ by Proposition 4.5. Lemma 7.2 shows how an $\hat{X}$ solving this new $\mathcal{H}$-twisted conjugacy problem leads to an X solving the earlier one, and that if $\hat{X}$ has length at most a constant times $|u|_H + |v|_H$, then the same will be true of X. We reapply Proposition 7.1, and again, if we hit case ($\mathcal{X}$1) or ($\mathcal{X}$2) then we can stop. In case ($\mathcal{X}$3), a short conjugator is found via Lemma 8.4. If we hit case ($\mathcal{X}$4) then we iterate down to a lower rank again.

The maximum number of times we can iterate through case ($\mathcal{X}$4) is $m-1$ times. This will bring us to rank 1 (if the process has not yet terminated) and then case ($\mathcal{X}$1) will apply. So this process will terminate at an X and so a W of suitably bounded length.

10. The algorithm: completing our proof of Theorem 2

Here, in outline, is our algorithm for the conjugacy and conjugacy search problems for H.

Input: Words u and v on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}, s^{\pm 1}$.

Step 1. Convert u and v to normal forms $\tilde{u} s^p$ and $\tilde{v} s^q$ respectively. If $p \neq q$, then stop and declare u is not conjugate to v. If $p = q < 0$, then replace u and v by their inverses and return to the start.

Time required: polynomial in $\ell(u) + \ell(v)$.

Step 2. If $p = q = 0$, then run the algorithm of Proposition 5.1 (III) solving the 0-twisted conjugacy problem. If it declares the 0-twisted conjugacy problem has no solution, then declare u is not conjugate to v. Otherwise it outputs a solution $(r, \tilde{w})$, so stop and declare $\tilde{w} s^r$ is a conjugator.

Time required: polynomial in $\ell(\tilde{u}) + \ell(\tilde{v})$.

Step 3. We have $p = q > 0$. Let $\mathcal{I}$ be the set of all pairs $(\tilde{w}, r)$, where $0 \leq r < p$, and $\tilde{w}$ is a word of the form UV where U is a prefix of $\tilde{u}$ and V^{-1} is a prefix of $\varphi^{-r}(\tilde{v})$. For each $(\tilde{w}, r)$ in $\mathcal{I}$, check whether $\tilde{u}\varphi^{-p}(\tilde{w}) = \tilde{w}\varphi^{-r}(\tilde{v})$ (as per the I-twisted conjugacy problem).

If a solution $(\tilde{w}, r)$ is found, then the algorithms declares that u and v are conjugate and outputs $\tilde{w}s^r$ as a conjugator. If no solution is found we continue to the next step.

Time required: the number of entries on the list $\mathcal{I}$ is bounded by a polynomial in $p + \ell(\tilde{u}) + \ell(\tilde{v})$ and the obvious solution to the word problem in F runs in linear time, so overall this step runs in time polynomial in $|p| + \ell(\tilde{v}) + \ell(\tilde{v})$.

Step 4. Run the algorithm of Corollary 8.3 for the $\mathcal{H}$-twisted conjugacy problem. If it declares there is no solution, stop and declare that u and v are not conjugate. If it exhibits a solution $(r, x, u_0, v_0, u_1, v_1)$, then declares that u and v are conjugate and output $u_0 x v_0^{-1} s^r$ as a conjugator.

Time required: polynomial in $|p| + |\tilde{u}|_H + |\tilde{v}|_H$.

Overall time required: Since $|\tilde{u}|_H \leq \ell(\tilde{u}) \leq C\ell(u)^m$ and $|\tilde{v}|_H \leq \ell(\tilde{v}) \leq C\ell(v)^m$ for a suitable constant $C > 0$, and $|p| \leq \ell(u)$, the total running time of the algorithm is polynomial in $\ell(u) + \ell(v)$.

Remark 10.1. The algorithm described above can be modified as follows to output in polynomial time a conjugator W (if one exists) with $\ell(W)$ at most a constant times $\ell(u) + \ell(v)$. We describe the required changes. First, when the algorithm of Lemma 8.2 finds a conjugator of form ($\mathcal{X}3$) its output includes the subword M_1 or M_1^{-1}. Add an extra step that replaces this M_1 by s^{pq}. (Lemma 8.4 confirms that the result remains a conjugator.) This will produce a word W_0 which is a conjugator whose *length in H* is bounded by a linear function of $|u|_H + |v|_H \leq \ell(u) + \ell(v)$. This means that there is a word W on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}, s^{\pm 1}$ that equals W_0 in H and has length $\ell(W) \leq \ell(u) + \ell(v)$. It remains to argue that we can further adapt the algorithm to exhibit such a W. The word W_0 is assembled from words derived from subwords of $\tilde{u}$ and $\tilde{v}$ as described in Section 9. We used Proposition 4.5 to bound the lengths (in H) of such subwords in terms of $|u|_H$ or $|v|_H$. Our proof of Proposition 4.5 is constructive. In particular it can be adapted to a polynomial time algorithm that, for example, takes a subword u' of $\tilde{u}$, where $u = \tilde{u}s^p$ in normal form, and gives a word on $a_1^{\pm 1}, \ldots, a_m^{\pm 1}, s^{\pm 1}$ that equals u' in H and whose length is at most a constant times $|u|_H$. We can then assemble W from words obtained in this manner.

11. An alternative approach

In this section we outline an alternative proof of Theorem 2 that is based on the structure of H_m as an iterated HNN extension. This alternative approach will be developed in detail in [BRSa] and applied to a wider class of free-by-cyclic groups.

We regard $H = H_m$ as an $(m-1)$-fold iterated HNN extension of $H_1 = \langle s, a_1 \rangle \cong \mathbb{Z}^2$ where at the j-th stage the base group is $H_j := \langle s, a_1, \ldots, a_j \rangle$, the stable letter is a_{j+1}, the associated (cyclic) subgroups are $\langle s \rangle$ and $\langle s a_j^{-1} \rangle$, and the relation $a_{j+1}^{-1} s a_{j+1} = s a_j^{-1}$ holds. This point of view enables one to argue by induction on m and appeal to the technology of corridors to analyse van Kampen diagrams and their annular analogues over the natural presentations of these groups. But in keeping with the viewpoint of this article, we shall suppress the use of diagrams here and concentrate on the algebraic translation of the insights that they provide.

There is a classical approach to the conjugacy problem in HNN extensions based on Collins' Lemma [Col69]—see [LS07], page 185, for example. This simplifies in the case where the associated subgroups are cyclic, as we shall now explain.

Let $G = (G_0, t \mid t^{-1}\alpha t = \beta)$ be an HNN extension where $A = \langle\alpha\rangle$ and $B = \langle\beta\rangle$ are infinite cyclic. We fix a generating set S for G_0 that includes α and β. A word U in the alphabet $S^{\pm 1}$ is in *cyclically reduced HNN form* if U^2 does not contain a pinch—i.e., a subword $t^{-1}ct$ with $c \in A$ or tdt^{-1} with $d \in B$.

For simplicity, we assume that distinct powers of β are not conjugate in G and that no power of α is conjugate to a power of β in G_0. A straightforward analysis of annular diagrams yields the following version of Collins' Lemma in this simplified setting.

Lemma 11.1 (Collins' Lemma)**.** *Assume that U and V are words in cyclically reduced HNN form. If U is conjugate to V in G, then either*

i U and V contain no occurrences of $t^{\pm 1}$ and either they are conjugate in G_0 or else one is conjugate into A and the other is conjugate into B; or else
ii both U and V contain an occurrence of $t^{\pm 1}$ and there are cyclic permutations U' of U and V' of V and an integer q such that $\alpha^{-q}U'\alpha^q = V'$ in G.

11.1. **The algorithm for Theorem 2.** We regard H_m as an HNN extension of H_{m-1} as described in the second paragraph. In the language used above, $G = H_m$ while $G_0 = H_{m-1}$, $A = \langle s\rangle$, $B = \langle\beta\rangle$ and $t = a_k$, where β is a generator we have added with $\beta = sa_{m-1}^{-1}$ in H_m.

Proceeding by induction, we may assume that we have a polynomial time algorithm to decide conjugacy in H_{m-1}. Given two words u, v in the generators of H_m (with β included) we rewrite them into cyclically HNN reduced words U, V. This is achieved by first transforming u and v to reduced HNN form by removing pinches and then examining cyclic permutations of u and v, removing any additional pinches that appear. The second step may need to be repeated several times, but the word is shortened each time. Both steps can be done in polynomial time without increasing the length of the words.

We are now able to apply Collins' Lemma. If there are no occurrences of $a_m^{\pm 1}$ in U' and V', then we are in case (i) and we apply the algorithm for H_{m-1}. Otherwise we are in case (ii) and we are left to determine if there is an integer p such that $s^{-q}U's^q = V'$. (Recall that $A = \langle\alpha\rangle = \langle s\rangle$.)

In polynomial time, we can rewrite U' and V' into normal form $\tilde{U}'s^r$ and $\tilde{V}'s^{r'}$, where the lengths of $\tilde{U}'$ and $\tilde{V}'$ are bounded polynomially by $|U|$ and $|V|$. If $r \neq r'$, then we stop and declare that U is not conjugate to V. If $r = r'$, then we are reduced to deciding if there is a positive integer p such that $\phi^p(\tilde{U}') = \tilde{V}'$ or $\phi^p(\tilde{V}') = \tilde{U}'$. The range of possible p is bounded by a linear function of $|U| + |V|$, by considerations of growth, as in Section 4. And for each specific p, we can evaluate $\phi^p(U)$ naively (letter by letter) and freely reduce to see if it is equal to V. As φ has polynomial growth, these evaluations can be done in polynomial time.

This algorithm, as we have described it, does not provide the linear upper bound on conjugator length that is required for Theorem 1. The main argument in [BRSa] overcomes this limitation with an alternative endgame that makes greater use of the structure of H_m as an iterated HNN extension.

References

[AS16] Y. Antolín and A. W. Sale. Permute and conjugate: the conjugacy problem in relatively hyperbolic groups. *Bull. Lond. Math. Soc.*, 48(4):657–675, 2016.

[BBMS97] G. Baumslag, M. R. Bridson, C. F. Miller, and H. Short. Finitely presented subgroups of automatic groups and their isoperimetric functions. *J. London Math. Soc. (2)*, 56(2):292–304, 1997.

[BFH00] M. Bestvina, M. Feighn, and M. Handel. The Tits alternative for Out(F_n). I. Dynamics of exponentially-growing automorphisms. *Ann. of Math. (2)*, 151(2):517–623, 2000.

[BFH05] M. Bestvina, M. Feighn, and M. Handel. The Tits alternative for Out(F_n). II. A Kolchin type theorem. *Ann. of Math. (2)*, 161(1):1–59, 2005.

[BFW19] M. Bestvina, K. Fujiwara, and D. Wigglesworth. The Farrell–Jones conjecture for free-by-cyclic groups. arViv:1906.00069, 2019.

[BG10] M. R. Bridson and D. Groves. The quadratic isoperimetric inequality for mapping tori of free group automorphisms. *Mem. Amer. Math. Soc.*, 203(955):xii+152, 2010.

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

[BH99] M. R. Bridson and A. Haefliger. *Metric Spaces of Non-positive Curvature*. Number 319 in Grundlehren der mathematischen Wissenschaften. Springer Verlag, 1999.

[BKS87] R. G. Burns, A. Karrass, and D. Solitar. A note on groups with separable finitely generated subgroups. *Bull. Austral. Math. Soc.*, 36(1):153–160, 1987.

[BMMV06] O. Bogopolski, A. Martino, O. Maslakova, and E. Ventura. The conjugacy problem is solvable in free-by-cyclic groups. *Bull. London Math. Soc.*, 38(5):787–794, 2006.

[BR] M. R. Bridson and T. R. Riley. The lengths of conjugators in the model filiform groups. In preparation.

[BR09] M. R. Bridson and T. R. Riley. Extrinsic versus intrinsic diameter for Riemannian filling-discs and van Kampen diagrams. *J. Diff. Geom.*, 82(1):115–154, 2009.

[BR20] O. Baker and T. Riley. Cannon-Thurston maps, subgroup distortion, and hyperbolic hydra. *Groups Geom. Dyn.*, 14(1):255–282, 2020.

[Bri02] M. R. Bridson. Polynomial Dehn functions and the length of asynchronously automatic structures. *Proc. London Math. Soc.*, 85(2):441–465, 2002.

[BRSa] M. R. Bridson, T. R. Riley, and A. Sale. Bounded conjugators in mapping tori of triangular free-group automorphisms. In preparation.

[BRSb] M. R. Bridson, T. R. Riley, and A. Sale. Conjugator length in finitely presented groups. In preparation.

[Bum15] I. Bumagin. Time complexity of the conjugacy problem in relatively hyperbolic groups. *Internat. J. Algebra Comput.*, 25(5):689–723, 2015.

[But07] J. O. Button. Mapping tori with first Betti number at least two. *J. Math. Soc. Japan*, 59(2):351–370, 2007.

[CM11] C. H. Cashen and N. Macura. Line patterns in free groups. *Geom. Topol.*, 15(3):1419–1475, 2011.

[Col69] D. J. Collins. On embedding groups and the conjugacy problem. *J. London Math. Soc. (2)*, 1:674–682, 1969.

[DER18] W. Dison, E. Einstein, and T. R. Riley. Taming the hydra: the word problem and extreme integer compression. *Internat. J. Algebra Comput.*, 28(7):1299–1381, 2018.

[DR13] W. Dison and T. R. Riley. Hydra groups. *Comment. Math. Helv.*, 88(3):507–540, 2013.

[EH06] D. Epstein and D. Holt. The linearity of the conjugacy problem in word-hyperbolic groups. *Internat. J. Algebra Comput.*, 16(2):287–305, 2006.

[Ger94] S. M. Gersten. Quadratic divergence of geodesics in CAT(0) spaces. *Geom. Funct. Anal.*, 4(1):37–51, 1994.

[JOR10] R. Ji, C. Ogle, and B. Ramsey. Relatively hyperbolic groups, rapid decay algebras and a generalization of the Bass conjecture. *J. Noncommut. Geom.*, 4(1):83–124, 2010. With an appendix by Ogle.

[Lev09] G. Levitt. Counting growth types of automorphisms of free groups. *Geom. Funct. Anal.*, 19(4):1119–1146, 2009.

[LS07] M. Lohrey and S. Schleimer. Efficient computation in groups via compression. In *Proc. Computer Science in Russia (CSR 2007)*, volume 4649 of *Lecture Notes in Computer Science*, pages 249–258. Springer, 2007.

[Lys89] I. G. Lysënok. Some algorithmic properties of hyperbolic groups. *Izv. Akad. Nauk SSSR Ser. Mat.*, 53(4):814–832, 912, 1989.

[Mac02] N. Macura. Detour functions and quasi-isometries. *Q. J. Math.*, 53(2):207–239, 2002.

[Mac13] N. Macura. CAT(0) spaces with polynomial divergence of geodesics. *Geom. Dedicata*, 163:361–378, 2013.

[O'C] Z. O'Connor. Conjugacy search problem for relatively hyperbolic groups. arXiv:1211.5561.

[Pue16] K. Pueschel. Hydra group doubles are not residually finite. *Groups Complex. Cryptol.*, 8(2):163–170, 2016.

[Sam06] P. Samuelson. On CAT(0) structures for free–by–cyclic groups. *Topology Appl.*, 153(15):2823–2833, 2006.

[Sch08] S. Schleimer. Polynomial-time word problems. *Comment. Math. Helv.*, 83(4):741–765, 2008.

Martin R. Bridson
Mathematical Institute, Andrew Wiles Building, Oxford OX2 6GG, United Kingdom
bridson@maths.ox.ac.uk, https://people.maths.ox.ac.uk/bridson/

Timothy R. Riley
Department of Mathematics, 310 Malott Hall, Cornell University, Ithaca, NY 14853, USA
tim.riley@math.cornell.edu, http://www.math.cornell.edu/~riley/

Andrew W. Sale
Department of Mathematics, University of Hawaii at Manoa, Honolulu, HI 96822, USA
http://math.hawaii.edu/~andrew/