Normal forms for automorphisms of universal Coxeter groups and palindromic automorphisms of free groups

Adam Piggott and Kim Ruane

2010

Abstract

We explicitly construct Markov languages of normal forms for the groups in the title of the paper and closely related groups. A Markov language of normal forms is a choice of "preferred spelling" for each group element such that the collection of choices is particularly simple in a language theoretic sense.

1 Introduction

It is a central difficulty of Combinatorial Group Theory that, given a finitely generated group G and a finite generating set $S \subset G$ which is closed under inverses, each element of G may be written as a product of generators in infinitely many ways. One might respond to this difficulty by making a choice of "preferred spelling" for each element of G; that is, one defines a section $N: G \to S^*$ of the natural projection $S^* \to G$, where S^* denotes the free monoid on S. The image $\mathcal{L} = N(G)$ of such a section is a called a *language of normal forms (in the alphabet S)*, and the image $N(g)$ of an element $g \in G$ is called *the normal form for g*. A well-chosen language of normal forms can allow one to perform calculations efficiently, or provide convenient additional structure with which to construct arguments. To realize such benefits one would generally like an efficient procedure for determining the normal form of an arbitrary element in G, and one would like to discover properties of the language $\mathcal{L}$ which inform about the group G. The theory of automatic groups, and related concepts, may be seen as a contemporary example of this approach [3].

In the vocabulary of [6], a *Markov language* is a subset of a free monoid that can be completely described by listing those subwords of length two which are 'permitted', and therefore also those that are 'forbidden'; that is, a word $s_1 \dots s_j$ is in the language if and only if each subword $s_i s_{i+1}$ is in the language. A Markov language of normal forms is necessarily regular, has an associated 'departure function' and comes equipped with a natural finite rewriting procedure, a procedure which aims to find the normal form corresponding to an arbitrary element $\mathsf{U} \in S^*$ but which, in general, may or may not terminate in a finite number of steps for a given input U (see Section 2 below for details). To say that a finitely generated group admits a Markov language of normal forms does not seem a particularly restrictive statement, but to exhibit such a language is a natural way to approach the study of a group. Some of the most familiar languages of normal forms are Markov, and the corresponding natural finite rewriting procedures are the essential part of well-known solutions to the word problem in these groups. For example, the standard languages of normal forms for finitely-generated free abelian groups, free groups and free products of finite groups are Markov languages, as are some normal forms associated to biautomatic structures on the braid groups and, more generally, Artin groups of finite type [1, Proposition 2.1].

We now introduce the groups which are studied in the present article. For each positive integer n, the *universal Coxeter group of rank* n is the group W_n presented by $\langle a_1, a_2, \ldots, a_n \mid a_1^2, a_2^2, \ldots, a_n^2 \rangle$. It is the simplest example of a free product of n groups, and one of the simplest examples of a Coxeter group. We write $\operatorname{Aut} \mathrm{W}_n$ for the group of automorphisms of W_n. The group of *outer automorphisms of* W_n, denoted $\operatorname{Out} \mathrm{W}_n$, is the quotient of $\operatorname{Aut} \mathrm{W}_n$ by the group of inner automorphisms $\operatorname{Inn} \mathrm{W}_n$. The group of *basis-conjugating automorphisms of* W_n, sometimes called the group of *pure symmetric automorphisms*, is the group $\operatorname{Aut}^0 \mathrm{W}_n$ of automorphisms which map each generator to a conjugate of itself. We write F_{n-1} for the free group of rank $n-1$. The group of *palindromic automorphisms of* F_{n-1}, denoted $\Pi\mathrm{F}_{n-1}$, is the group of automorphisms which map each element of a fixed basis to an element which reads the same forwards and backwards [2]. Throughout this article we shall assume that $n \geq 3$, since the cases $n \leq 2$ are atypical and are easily dealt with individually.

The automorphism groups introduced in the paragraph above are closely related, as explained in Section 4 below. Understanding $\operatorname{Aut}^0 \mathrm{W}_n$ is key to understanding $\operatorname{Aut} \mathrm{W}_n$, $\operatorname{Out} \mathrm{W}_n$ and $\Pi\mathrm{F}_{n-1}$. We are particularly interested in $\operatorname{Aut} \mathrm{W}_n$ for the purposes of comparing its properties to $\operatorname{Aut} \mathrm{F}_{n-1}$, noting that $\operatorname{Aut} \mathrm{W}_n$ embeds in $\operatorname{Aut} \mathrm{F}_{n-1}$ and, in particular, $\operatorname{Aut} \mathrm{W}_3$ is isomorphic to $\operatorname{Aut} \mathrm{F}_2$ (see, for example, [10, Remark 2]).

In the present article we explicitly construct Markov languages of normal forms for $\operatorname{Aut}^0 \mathrm{W}_n$, $\operatorname{Aut} \mathrm{W}_n$ and $\Pi\mathrm{F}_{n-1}$, and a regular language of normal forms for $\operatorname{Out}^0 \mathrm{W}_n$. These languages are not languages of geodesics and they are not part of automatic structures for the corresponding groups. Our language for $\operatorname{Aut}^0 \mathrm{W}_3$ is part of an asynchronously automatic structure (see Remark 5.4 below), but it is not an optimal choice of normal form for the group. That particular group is biautomatic—one may invoke [3, Theorem 4.1.4] because $\operatorname{Aut}^0 \mathrm{W}_3$ has finite index in $\operatorname{Aut} \mathrm{W}_3$ and $\operatorname{Aut} \mathrm{W}_3$ is biautomatic [10]—and a biautomatic structure is superior to an asynchronously automatic structure in a number of ways. We were unable to determine whether or not the other languages constructed in this paper are part of asynchronously automatic structures, and indeed it remains unknown whether or not the associated groups admit any type of automatic structure.

This article was inspired by [6], in which the authors exhibit a regular language of normal forms for $\mathrm{Aut}^0 \mathrm{F}_n$, the basis-conjugating automorphisms of F_n. The strategy of proof is to find an infinite generating set $T \subset \mathrm{Aut}^0 \mathrm{F}_n$ (denoted S in [6]) and an 'initial function' $I\colon \mathrm{Aut}^0 \mathrm{F}_n \to T$ which satisfies three properties (Properties (I1), (I2) and (I3) in Section 3.2). One obtains a section $N\colon \mathrm{Aut}^0 \mathrm{F}_n \to T^*$ of the evaluation map $T^* \to \mathrm{Aut}^0 \mathrm{F}_n$ from the first two properties. The third property implies that the image $N(\mathrm{Aut}^0 \mathrm{F}_n)$ satisfies a definition analogous to our definition of a Markov language, but with T infinite. The authors then define a finite generating set $T_0 \subset T$ and construct a map $N(\mathrm{Aut}^0 \mathrm{F}_n) \to T_0^*$. The result is a regular, but not Markov, language of normal forms for $\mathrm{Aut}^0 \mathrm{F}_n$. This recipe for cooking up a subset of T^* which satisfies the Markov property is powerful, and the technicalities of applying it to the basis-conjugating automorphisms of F_n adapt neatly to the basis-conjugating automorphisms of W_n. In particular, we are able to do so using a finite generating set $S_n \subset \mathrm{Aut}^0 \mathrm{W}_n$ and hence find a Markov language of normal forms.

In addition to its clear connection to [6], the vocabulary and ideas in the present article also tie in with other relatively recent contributions to the theory of the automorphisms of free products of groups (see, for example, [8] and the references therein). These contributions have roots in the work of McCullough and Miller [9] (the 'MM' in MM-trees), who constructed a contractible space on which the group of 'symmetric automorphisms' of a free product of groups acts with a strict fundamental domain.

We now describe the structure of the present article. Section 2 contains background material on Markov languages and rewriting systems. In Section 3 we define a generating set $S \subset \mathrm{Aut}^0 \mathrm{W}_n$ and an initial function which determines a Markov language of normal forms for $\mathrm{Aut}^0 \mathrm{W}_n$. We then construct related languages of normal forms for $\mathrm{Aut}\, \mathrm{W}_n$, $\Pi\mathrm{F}_{n-1}$ and $\mathrm{Out}\, \mathrm{W}_n$ in Section 4. In Section 5 we establish some properties of our language of normal forms for $\mathrm{Aut}^0 \mathrm{W}_n$. Section 6 contains the most technical part of our proof that the function defined in Section 3 is indeed an initial function.

2 Markov languages and rewriting systems

Throughout this section G will denote a finitely generated group, and $S \subset G$ a finite generating set which is closed under inverses and does not include the identity element.

A (deterministic) finite state automaton over the alphabet S is a finite directed graph in which vertices are called *states*; directed edges are called *transitions*; each state is either an *accept state*, a *failure state* or an *intermediate state*; one state is further identified as the *initial state*; each transition is labeled with an element of S; and there is exactly one transition leaving v with label s for each state v and each generator $s \in S$. The language recognized by a finite state automaton is the subset $\mathcal{L} \subset S^*$ of labels of paths which start at the initial state and end at an accept state. A *regular language* is a language which is recognized by a finite state automaton. Groups which admit a regular language of normal forms were first studied by Gilman [5], and have since enjoyed a prominent role in geometric group theory because of their role in the theory of automatic groups, as mentioned in the following paragraph.

For a word $\mathsf{U} \in S^*$ and a non-negative integer t, we write $\mathsf{U}(t)$ for the prefix of U that has t letters, interpreting $\mathsf{U}(t)$ as U in the case that U has less than t letters. A language of normal forms $\mathcal{L} \subset S^*$ is said to have the *fellow-traveler property* if there exists a constant K, called the *fellow-traveler constant*, such that following is true: for each pair $\mathsf{U}, \mathsf{V} \in \mathcal{L}$ such that the group element $\mathsf{U}^{-1}\mathsf{V}$ is in S, and for each non-negative integer t, the group element $\mathsf{U}(t)^{-1}\mathsf{V}(t)$ can be spelled in K letters or less. This means that $\mathcal{L}$ has the fellow-traveler property if two particles remain uniformly close whenever they travel at unit speed along paths in $\Gamma_S(G)$ which correspond to normal forms that end at adjacent vertices. The *asynchronous fellow-traveler property* is a strictly weaker property, where one is allowed to vary the speed of the particles in order to ensure that they stay uniformly close. A *departure function* for $\mathcal{L}$ is a function $D: \mathbb{N} \to \mathbb{N}$ such that, for all words $\mathsf{U} \in S^*$ such that U is a subword of some word $\mathsf{W} \in \mathcal{L}$, if U has at least $D(r)$ letters, then the group element represented by U cannot be spelled in r letters or less. We say that $\mathcal{L}$ is part of an *automatic structure* on G if $\mathcal{L}$ is regular and has the fellow-traveler property, and part of an *asynchronously automatic structure* on G if $\mathcal{L}$ is regular, has the asynchronous fellow-traveler property and there exists a departure function for $\mathcal{L}$. Such structures impose geometric and algorithmic restrictions on G. The terms introduced in this paragraph are carefully developed and explained in [3].

Recall that a Markov language in the alphabet S is a language $\mathcal{L} \subset S^*$ which is defined by identifying each word of length two as either permitted or forbidden. We note that a Markov language necessarily contains the empty word, denoted 1, and all words of length one. A Markov language is regular because it is recognized by the finite state automaton with the following simple structure: there are $|S|+1$ accept states and one failure state; one of the accept states is designated as the initial state and the remainder are each labeled by a different element of S; there is a transition labeled s from the initial state to the state s for each $s \in S$; there is a transition labeled s from the state t to the state s if and only if ts is permitted; there is a transition labeled s from the state t to the failure state if and only if ts is forbidden; there is a transition labeled s from the failure state to the failure state for each $s \in S$. A Markov language has a departure function because if W is in $\mathcal{L}$, then so is each subword U of W, and one may define $D(r)$ to exceed the maximum length of a normal form for those group elements spelled by r letters or less.

A *finite rewriting system (over the alphabet S)* is a finite set of ordered pairs $(\mathsf{X}, \mathsf{Y}) \in S^* \times S^*$. A word $\mathsf{U} \in S^*$ is transformed into the word $\mathsf{V} \in S^*$ by *application* of the rule (X, Y) if V is obtained from U by replacing an occurrence of the subword X by the subword Y. A *finite rewriting system for G* is a finite rewriting system such that words $\mathsf{U}, \mathsf{V} \in S^*$ spell the same element of G if and only if U can be obtained from V by application of rewriting rules.

There is a natural finite rewriting system associated to a Markov language of normal forms $\mathcal{L}$. Each rewriting rule involves replacing a forbidden subword X by the normal form for the group element spelled by X. From this we define a rewriting procedure: given a word which is not in normal form, replace the left-most forbidden subword by the corresponding normal form; repeat this until there are no forbidden subwords in the result. We shall refer to this procedure as the *rewriting procedure associated to* $\mathcal{L}$, and each replacement of a forbidden subword is called a *(rewriting) step.* One hopes to discover that this rewriting procedure terminates in a finite number of steps for an arbitrary input, in which case we say that the rewriting procedure is *complete* and we have a natural algorithm for finding the normal form of a group element. One often shows that a rewriting procedure is complete by finding a total order $>$ on S^* with the following property: if the result of a rewriting step is that $\mathsf{U} \in S^*$ is transformed to $\mathsf{V} \in S^*$, then $\mathsf{U} > \mathsf{V}$. Because we have defined our procedure to always replace the left-most forbidden pair, an induction on the number of letters to the right of the left-most forbidden pair can be used to show that our rewriting procedure is complete if and only if it terminates in a finite number of steps for an arbitrary input of the form $s_1 s_2 \dots s_j$, where $s_{j-1}s_j$ is the only forbidden pair.

Example 2.1. The language of reduced words in the alphabet $A_n := \{a_1, \dots, a_n\}$ is the language $\mathcal{R}_n \subset A_n^*$ in which words of the form $a_i a_i$ are forbidden. Clearly, $\mathcal{R}_n$ is a Markov language of normal forms for the group W_n. Each rewriting rule in the associated rewriting system involves replacing a forbidden word by the empty word. It is clear that if $\mathsf{U} \in A_n^*$ is transformed to $\mathsf{V} \in A_n^*$ by application of such a rule, then $\mathsf{U} > \mathsf{V}$ in the 'shortlex' order on A_n^* (where $\mathsf{X} > \mathsf{Y}$ in the shortlex order if X is longer than Y, or X has the same length as Y but X follows Y in the dictionary order determined by some fixed order on A_n). Thus the rewriting procedure associated to $\mathcal{R}_n$ is complete and provides a solution to the word problem in W_n. For this particular example we need not insist that one always replaces the left-most forbidden subword in a rewriting step, as any replacement will result in a shorter word. In such a case it is usual to say that the *rewriting system is complete.*

We shall write $\mathcal{M}$ for the family of finitely generated groups which admit a Markov language of normal forms. This family is 'closed' in the following ways.

Lemma 2.2. *(a) $\mathcal{M}$ is closed under the operation of free product;*

(b) $\mathcal{M}$ is closed under extension in the following sense: if there exists a short exact sequence $1 \to N \to E \to Q \to 1$ and N and Q are in $\mathcal{M}$, then E is also in $\mathcal{M}$;

(c) if H has finite index in G and H is in $\mathcal{M}$, then G is also in $\mathcal{M}$.

Each of the above statements has a constructive proof, so if one can decompose a group G into factors with known Markov languages of normal forms, and the decomposition uses the limited operations above to combine the factors, then one has effectively exhibited a Markov language of normal forms for G. For example, natural Markov languages of normal forms for finitely-generated free groups and nilpotent groups follow from their decompositions into finite groups and copies of the infinite cyclic group. It is more interesting to discover Markov languages of normal forms which do not arise in this way. For example, Garside exhibited Markov languages of normal forms for the braid groups, and Charney generalized this work to exhibit Markov languages of normal forms for Artin groups of finite-type [1].

It follows from "The Pumping Lemma" for regular languages [3, Lemma 1.2.13 and Example 2.5.12]) that infinite torsion groups do not admit Markov languages of normal forms.

It is immediate from the definition that a Markov language of normal forms $\mathcal{L}$ is prefix-closed; that is, if $\mathsf{U} \in \mathcal{L}$ then so is every initial subword of U (Markov languages are also factor-closed, a strictly stronger property). Note also that, since $1 \in \mathcal{L}$, each subword of the form ss^{-1} is forbidden (that is, each element of $\mathcal{L}$ is freely reduced). Thus $\mathcal{L}$ corresponds to a spanning subtree $\mathcal{T}$ of the Cayley graph $\Gamma_S(G)$. The fact that the language is Markov is saying something about the algorithmic complexity of building $\mathcal{T}$ since: $\mathcal{T}$ contains every edge adjacent to the identity vertex v_1; given a subtree $\mathcal{T}_0 \subset \mathcal{T}$ such that $\mathcal{T}_0$ contains the vertex v_1, and given a labeled edge E which is not in $\mathcal{T}_0$ but which has an endpoint $v \neq v_1$ in $\mathcal{T}_0$, then one can decide whether or not E is in $\mathcal{T}$ with only the 'local' information of the label on the last edge along the unique reduced path in $\mathcal{T}_0$ from v_1 to v.

3 Normal forms for $\mathrm{Aut}^0 \mathrm{W}$

In this section we focus on the group $\mathrm{Aut}^0 \mathrm{W}$. We define a generating set $S \subset \mathrm{Aut}^0 \mathrm{W}$, the set of simple automorphisms, and a function $I\colon \mathrm{Aut}^0 \mathrm{W} \to S \cup \{1\}$. We claim that I has the three properties, called (I1), (I2) and (I3), which make it an initial function. Then $I(f)$ is the first letter in the normal form for f, and thus I determines a language of normal forms $\mathcal{L}_I$ for $\mathrm{Aut}^0 \mathrm{W}$. The proof of the claim is technical, and is postponed to Section 6.

Notation 3.1. We now establish some notational conventions for the remainder of the paper. Our notation shall follow that of [6] as closely as possible. In particular, we shall write automorphisms acting on the right, so that a product of automorphisms is to be read from left to right. We shall once and for all fix an integer $n \geq 3$, and we write $\mathbf{n} := \{1, 2, \ldots, n\}$. Unless it will cause ambiguity, we shall omit the subscript n from variables, writing W for W_n, A for A_n, F for F_n etc.

3.1 Simple automorphisms

Recall that $\mathrm{Aut}^0 \mathrm{W} \subset \mathrm{Aut}\, \mathrm{W}$ consists of those automorphisms which map each generator a_i to a conjugate of itself. For distinct integers $i, k \in \mathbf{n}$, we write $x_{ik} \in \mathrm{Aut}^0 \mathrm{W}$ for the automorphism determined by $a_i \mapsto a_k a_i a_k$, and $a_j \mapsto a_j$ for $j \in \mathbf{n} \setminus \{i\}$. Such an automorphism is called a *partial conjugation.* Although the set of partial conjugations is a natural and well-known generating set for $\mathrm{Aut}^0 \mathrm{W}$ (see, for example, [4]), it does not suit our purposes. Here we define a larger generating set S, the set of simple automorphisms. The definition of

S which seems most natural makes reference to the Cayley graph for W with respect to A, so we first describe that. Our description of this Cayley graph is not standard, although it is equivalent to the standard description, and we have some unusual conventions for depicting its subgraphs. These unusual features serve to highlight the relationship between the ideas herein and those in [6] and [9].

Let $\Gamma_A(W)$, usually written Γ, be the regular undirected tree with valence n and edges of unit length. Label the edges of Γ by letters in the set A such that each vertex has exactly one incident edge with label a_i for each $i \in \mathbf{n}$. Fix a vertex v_1 of Γ to be viewed as the identity vertex. Then Γ is the *Cayley graph of* W *with respect to* A. For each $w \in \mathrm{W}$, write v_w for the endpoint of the path from v_1 with label w. We shall refer to the vertices of Γ as *white vertices*, the midpoints of edges as *black vertices* and the subset of Γ between a black vertex and an adjacent white vertex as a *half-edge*.

Remark 3.2. We describe our conventions for depicting Γ and subsets of Γ. Black vertices are indicated by filled circles and are labeled by the corresponding element of A. The white vertex v_1 is indicated by an unfilled star, and the remaining white vertices are indicated by unfilled circles. We omit the labels on white vertices, as the label on a white vertex v may be determined by reading off the labels on the black vertices along the unique geodesic from v_1 to v. In Fig. 1 we depict Γ in the case that $n = 3$.

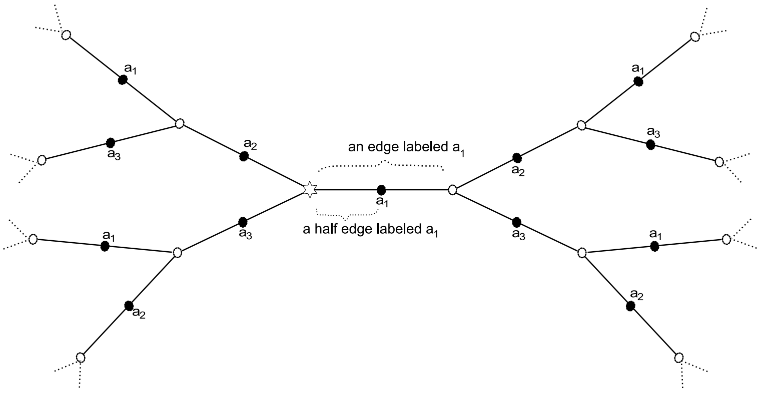

Figure 1: Γ in the case that $n = 3$.

Consider the natural left-action $\rho : \mathrm{W} \to \operatorname{Isom}\Gamma$ corresponding to left-multiplication; that is, $\rho(w_1)v_{w_2} = v_{w_1w_2}$ for $w_1, w_2 \in \mathrm{W}$. Let $t \in \mathrm{W}$ be an involution. Then t is of the form $wa_i\overline{w}$ for some $i \in \mathbf{n}$ and some $w \in \mathrm{W}$, where $\overline{w}$ denotes the inverse of w (which can be spelled by reversing a word for w). The isometry $\rho(t)$ is a reflection which fixes only the midpoint of the path from v_1 to v_t; we denote this fixed point F_t. Suppose $f \in \operatorname{Aut}^0 \mathrm{W}$ and $i \in \mathbf{n}$, so a_if is an involution. It follows from the definition of $\operatorname{Aut}^0 \mathrm{W}$ that each fixed point F_{a_if} is a black vertex contained in an edge labeled by a_i, and one may determine the image $a_i f = w_i a_i \overline{w_i}$ from the position of $F_{a_i f}$ by noting that the unique geodesic from v_1 to the edge containing $F_{a_i f}$ is labeled by w_i. Thus the minimal connected subspace $\Delta_f \subset \Gamma$ which contains the white vertex v_1 and the black vertices $F_{a_1 f}, \dots, F_{a_n f}$ encodes sufficient information to determine f (provided that the white vertex v_1 is clearly indicated, and each black vertex is labeled by an element of A or as one of the fixed points $F_{a_1 f}, \dots F_{a_n f}$). We call Δ_f the *automorphism subtree for f.*

Remark 3.3. When depicting subsets of Γ, it will be our convention to omit the label a_i from a black-vertex which is labeled $F_{a_i f}$ for some $f \in \operatorname{Aut}^0 \mathrm{W}$.

Example 3.4. In Fig. 2 we show the automorphism subtree Δ_g for $g = x_{21}x_{12}x_{21}x_{31} \in \operatorname{Aut}^0 \mathrm{W}_3$. Note that $a_1 g = a_1a_2a_1a_2a_1$, $a_2 g = a_1a_2a_1a_2a_1a_2a_1$ and $a_3 g = a_1a_3a_1$.

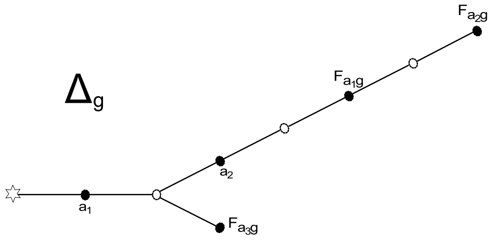

Figure 2: Δ_g for $g = x_{21}x_{12}x_{21}x_{31} \in \operatorname{Aut}^0 \mathrm{W}_3$.

Each automorphism subtree contains at least n black vertices. The number of black vertices in an automorphism subtree may be considered a measure of the complexity of the corresponding automorphism. Our preferred set of generators for $\operatorname{Aut}^0 \mathrm{W}$ consists of those non-identity automorphisms which are minimal in this sense.

Definition 3.5 (Simple automorphisms and MM-trees). Those automorphism subtrees which contain exactly n black vertices are called *(reduced labeled rooted) MM-trees.* An automorphism $f \in \operatorname{Aut}^0 \mathrm{W} \setminus \{\mathrm{id}\}$ is called a *simple automorphism* if the corresponding automorphism subtree is an MM-tree. The set of simple automorphisms is denoted by S. We write simple automorphisms as products of partial conjugations enclosed by square brackets; for example, $[x_{21}x_{31}x_{23}] \in S_3$.

Note that S includes the partial conjugations and, in contrast to [6], we exclude the identity automorphism from S. It is immediate from the definitions that S is a finite set. A connected subset $\Delta \subset \Gamma$ is an MM-tree if and only if it is a union of half-edges which contains the white vertex v_1 (and possibly other white vertices) and exactly n black vertices, which must then be labeled $F_{a_1 f}, \dots, F_{a_n f}$, and each white vertex, except perhaps v_1, has valence at least two. It also follows from the definitions that each element of S is an involution (this is most easily seen from the characterization in Remark 3.7 below).

Example 3.6. In Fig. 3 we show the MM-trees for the identity automorphism and the 15 elements of S_3.

Remark 3.7. Following [6], one may characterize the simple automorphisms as those non-identity elements of $\mathrm{Aut}^0\,\mathrm{W}$ which can be written in the form $t_1t_2\dots t_n$, where

$$t_k = \prod_{\ell\in P_k} x_{\ell k}$$

for some $P_k \subset \mathbf{n}\setminus\{k\}$ and t_i commutes with t_j for each pair $i,j\in\mathbf{n}$.

3.2 The language $\mathcal{L}_I$

The definition of a simple automorphism leads naturally to the following function. Recall that we write 1 for the empty word in S^*.

Definition 3.8 (The initial function I). Let $I\colon \mathrm{Aut}^0\,\mathrm{W} \to S\cup\{1\}$ be the function defined as follows: $I(\mathrm{id}) = 1$; for each $f\in \mathrm{Aut}^0\,\mathrm{W}\setminus\{\mathrm{id}\}$, $I(f)$ is the simple automorphism whose MM-tree is obtained from Δ_f by collapsing all edges except those containing the points $F_{a_1f},\dots,F_{a_nf}$.

Example 3.9. In Fig. 4 we show the automorphism subtrees for $g = x_{21}x_{12}x_{21}x_{31}\in \mathrm{Aut}\,\mathrm{W}_3$ and $I(g) = x_{21}$.

The following lemma establishes that the function I is well-defined.

Lemma 3.10. *Let $f\in\mathrm{Aut}^0\,\mathrm{W}$ and let Δ be the MM-tree obtained from Δ_f by collapsing all edges except those containing the points $F_{a_1f},\dots,F_{a_nf}$. Then $\Delta=\Delta_{\mathrm{id}}$ if and only if $f=\mathrm{id}$.*

Proof. Only one direction requires proof. Suppose that $\Delta=\Delta_{\mathrm{id}}$. Let $i,j\in\mathbf{n}$ be distinct integers. Since $\Delta=\Delta_{\mathrm{id}}$, the geodesic from v_1 to F_{a_if} does not pass through F_{a_jf}, and the geodesic from v_1 to F_{a_jf} does not pass through F_{a_if}. It follows that more than half of the reduced word for a_if, and more than half of the reduced word for a_jf, remain uncanceled when the concatenation $(a_if)(a_jf)$ is reduced to a reduced word. It follows in turn that, for each reduced word $a_{i_1}a_{i_2}\dots a_{i_k}$ in the alphabet A, at least one letter from each $a_{i_\ell}f$ remains uncanceled when one reduces the concatenation $(a_{i_1}f)(a_{i_2}f)\dots(a_{i_k}f)$ to a reduced word. So the reduced word corresponding to $(a_{i_1}f)(a_{i_2}f)\dots(a_{i_k}f)$ has at least k letters. Let $\ell\in\mathbf{n}$. Since f is surjective, there exists a reduced word $a_{\ell_1}a_{\ell_2}\dots a_{\ell_k}$ in the alphabet A such that $(a_{\ell_1}f)(a_{\ell_2}f)\dots(a_{\ell_k}f) = a_\ell$. It follows that $k=1$ and, since $f\in\mathrm{Aut}^0\,\mathrm{W}$, $a_{\ell_1}=a_\ell$. Thus $f=\mathrm{id}$. □

For each $w\in\mathrm{W}$, we write $|w|$ for the length of w with respect to the generating set A. For each $f\in\mathrm{Aut}^0\,\mathrm{W}$ and $i\in\mathbf{n}$, we write $w_{i,f}\in\mathrm{W}$ for the minimal length element such that $a_if = w_{i,f}a_i\overline{w_{i,f}}$, and we write

$$\|f\| := \sum_{i\in\mathbf{n}} |w_{i,f}|\,.$$

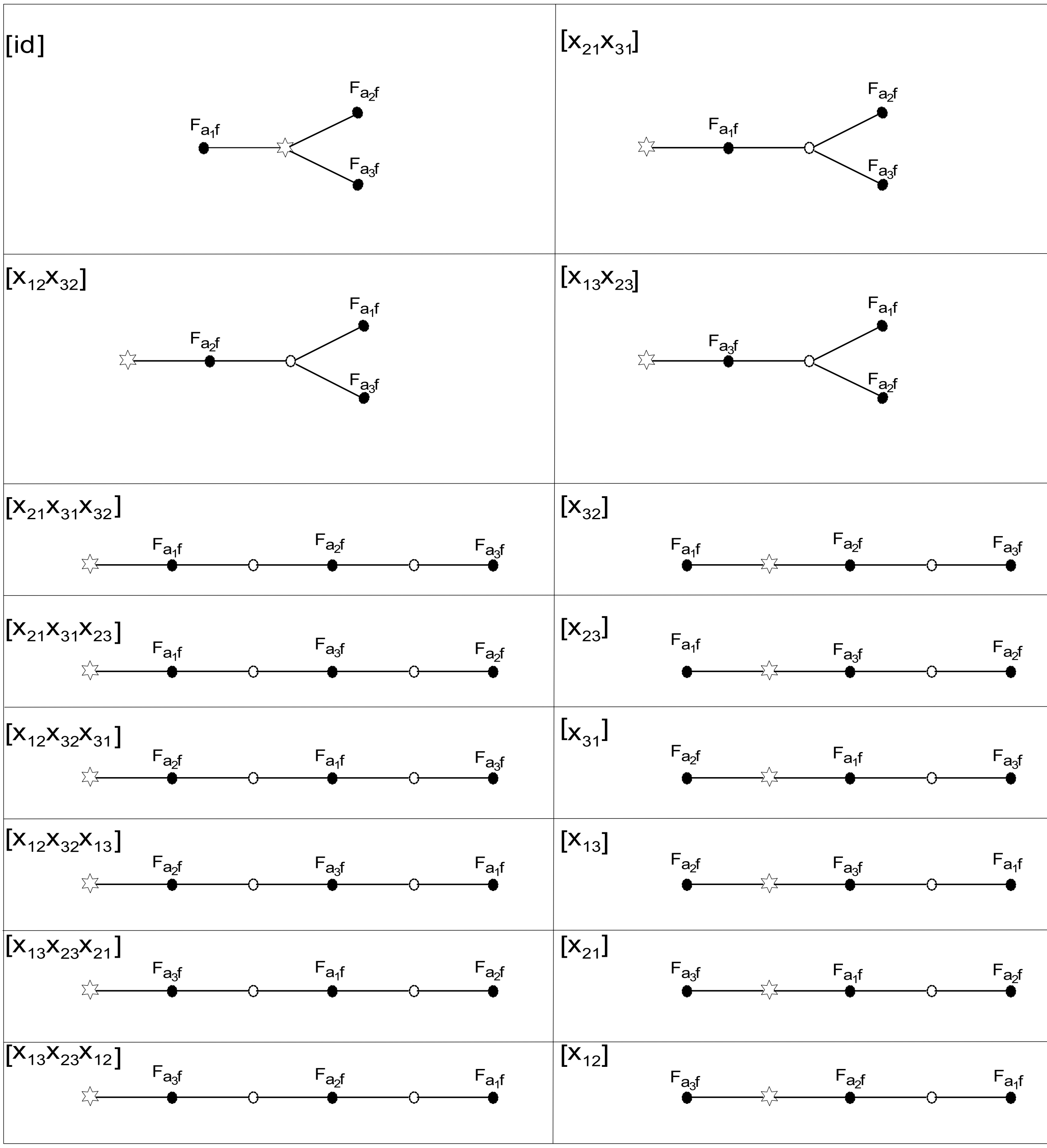

Figure 3: MM-trees for the identity automorphism and the elements of S_3.

Thus $\|f\|$ measures of the complexity of the automorphism f by summing the lengths of the "conjugating words" in the images of generators.

We claim that I satisfies the following properties for each $s \in S$ and $f \in \operatorname{Aut}^0 \mathrm{W} \setminus \{\mathrm{id}\}$:

(I1) $I(\mathrm{id}) = 1$, $I(s) = s$ and $I(f) \in S$;

(I2) $\|I(f)f\| < \|f\|$;

(I3) $I(sf) = s$ if and only if $I(sI(f)) = s$.

Property (I1) is immediate from the definitions and Lemma 3.10. In Section 6 we prove that Properties (I2) and (I3) hold.

In the vocabulary of [6], Properties (I1), (I2) and (I3) imply that I is an *initial function* (we remind the reader that each $s \in S$ is an involution, so we do not need to invert elements in the definition of (I2), as is necessary in [6]). Properties (I1) and (I2) imply that I determines a normal form function $N\colon \operatorname{Aut}^0 \mathrm{W} \to S^*$ as follows: $N(\mathrm{id}) = 1$; for $f \in \operatorname{Aut}^0 \mathrm{W} \setminus \{\mathrm{id}\}$, $N(f) = s_1 s_2 \dots s_m$ where

$$s_1 = I(f), s_2 = I(s_1 f), \dots, s_m = I(s_{m-1} \dots s_2 s_1 f) \text{ and } s_m \dots s_2 s_1 f = \mathrm{id}.$$

We write $\mathcal{L}_I$ for the image of N. The next lemma shows that, because of Property (I3), $\mathcal{L}_I$ is a Markov language of normal forms for $\operatorname{Aut}^0 \mathrm{W}$.

Lemma 3.11. *$\mathcal{L}_I$ is the Markov language in which a length-two word $st \in S^*$ is permitted if and only if $I(st) = s$.*

Proof. Let $\mathcal{L} \subset S^*$ denote the Markov language in which a length-two word $st \in S^*$ is permitted if and only if $I(st) = s$.

First we show that $\mathcal{L}_I \subset \mathcal{L}$. Let $s_1 s_2 \dots s_m \in \mathcal{L}_I$. Since $\mathcal{L}$ contains 1 and all words of length one, we may assume that $m \geq 2$. We must show that $I(s_i s_{i+1}) = s_i$ for each $1 \leq i \leq m-1$. It follows immediately from the definitions that $s_i = I(s_i s_{i+1} \dots s_m)$ for each $1 \leq i \leq m-1$. Then

$$s_i = I(s_i s_{i+1} \dots s_m) = I(s_i I(s_{i+1} \dots s_m)) = I(s_i s_{i+1}),$$

where the second equality follows from Property (I3).

Now we show that $\mathcal{L} \subset \mathcal{L}_I$. Let $s_1 s_2 \dots s_m \in \mathcal{L}$. Let f denote the automorphism spelled by $s_1 s_2 \dots s_m$. Property (I1) implies that $\mathcal{L}_I$ contains 1 and all words of length one, so we may assume that $m \geq 2$. It is immediate that $s_m \dots s_2 s_1 f = \mathrm{id}$. It remains to show that $s_k = I(s_{k-1} \dots s_1 f)$ for each $1 \leq k \leq m$. It is immediate that $s_m = I(s_m) = I(s_{m-1} \dots s_2 s_1 f)$. Assume that $s_k = I(s_k \dots s_m) = I(s_{k-1} \dots s_1 f)$ for some $2 \leq k \leq m$. Then

$$s_{k-1} = I(s_{k-1} s_k) = I(s_{k-1} s_k \dots s_m) = I(s_{k-2} \dots s_1 f),$$

where the first equality follows from the definition of $\mathcal{L}$, and the second from the inductive hypothesis together with Property (I3). By Induction, we have that $s_k = I(s_{k-1} \dots s_1 f)$ for each $1 \leq k \leq m$. $\square$

4 Normal forms for related groups

We now turn our attention to groups closely related to $\operatorname{Aut}^0 \mathrm{W}$.

First we consider the group $\operatorname{Aut}\mathrm{W}$. There are exactly n distinct conjugacy classes of involutions in W, one for each of the generators $a_1, \ldots, a_n$. An automorphism of W must permute the conjugacy classes of involutions, and each permutation of the set $\{a_1, \ldots, a_n\}$ determines an automorphism of W, thus we have a transitive action of $\operatorname{Aut}\mathrm{W}$ on a set of size n. The subgroup $\operatorname{Aut}^0\mathrm{W}$ is the kernel of this action, so we have a short exact sequence $1 \to \operatorname{Aut}^0\mathrm{W} \to \operatorname{Aut}\mathrm{W} \to \Sigma_n \to 1$, where Σ_n denotes the symmetric group on a set of size n. By Lemma 2.2(b) we have the following.

Corollary 4.1. *The concatenation of the language $\mathcal{L}_I$ and the language $\Sigma_n \setminus \{1\}$ is a Markov language of normal forms for* $\operatorname{Aut}\mathrm{W}$

We now exhibit a Markov language of normal forms for the palindromic automorphisms of a free group $\Pi\mathrm{F}_{n-1}$. Throughout this paragraph we will need to use subscripts to indicate the ranks of the various groups involved. The key observation here, made by Miller and recorded in [8, Lemma 6.1], is that $\Pi\mathrm{F}_{n-1}$ is isomorphic to $\operatorname{Aut}(\mathrm{W}_n, a_1)$, the subgroup of automorphisms which fix a_1. The isomorphism $\operatorname{Aut}(\mathrm{W}_n, a_1) \cong \Pi\mathrm{F}_{n-1}$ is seen as follows: the subset $\mathrm{E}_n \subset \mathrm{W}_n$ of even length elements is a characteristic and free subgroup which is freely generated by $B = \{a_1a_2, a_1a_2, \ldots, a_1a_n\}$; the centralizer of E_n in W_n is trivial, thus the homomorphism $\rho\colon \operatorname{Aut}\mathrm{W}_n \to \operatorname{Aut}\mathrm{E}_n$ determined by restriction is an injection [11]; the image of $\operatorname{Aut}(W, a_1)$ under ρ is precisely the palindromic automorphisms of E_n. Thus it suffices to consider $\operatorname{Aut}(\mathrm{W}_n, a_1)$. Let (S_n, a_1) denote the set of simple automorphisms which fix a_1; that is, $(S_n, a_1) := S_n \cap \operatorname{Aut}(\mathrm{W}_n, a_1)$. It follows from the definition of I that if $f \in \operatorname{Aut}^0(\mathrm{W}, a_1)$, then $I(f) \in (S_n, a_1) \cup \{\mathrm{id}\}$. It follows that the restriction $N|_{\operatorname{Aut}^0(\mathrm{W}, a_1)}$ determines a Markov language of normal forms $(\mathcal{L}_I, a_1)$ for $\operatorname{Aut}^0(\mathrm{W}_n, a_1) = \langle (S_n, a_1) \rangle = \operatorname{Aut}^0\mathrm{W}_n \cap \operatorname{Aut}(\mathrm{W}_n, a_1)$. By an argument analogous to that used to establish the corollary above, we have a short exact sequence $1 \to \operatorname{Aut}^0(\mathrm{W}_n, a_1) \to \operatorname{Aut}(\mathrm{W}_n, a_1) \to (\Sigma_n, a_1) \cong \Sigma_{n-1} \to 1$. By Lemma 2.2(b) we have the following.

Corollary 4.2. *The concatenation of the language $(\mathcal{L}_I, a_1)$ and the language $\Sigma_{n-1} \setminus \{1\}$ is a Markov language of normal forms for* $\operatorname{Aut}(W_n, a_1) \cong \Pi\mathrm{F}_{n-1}$.

We now consider the outer automorphisms of W. The group $\operatorname{Out}\mathrm{W}$ acts transitively on the set of conjugacy classes of involutions in W. We write $\operatorname{Out}^0\mathrm{W}$ for the kernel of this action, so we have the short exact sequence $1 \to \operatorname{Out}^0\mathrm{W} \to \operatorname{Out}\mathrm{W} \to \Sigma_n \to 1$. It is shown in [7] that $\operatorname{Out}^0\mathrm{W}$ is isomorphic to the subgroup of $\operatorname{Aut}^0\mathrm{W}$ generated by the following set of partial conjugation:

$$\{x_{ik} \mid 1 \leq i, k \leq n, i \neq 1, i \neq k, (i, k) \neq (2, 1)\}.$$

We shall identify $\operatorname{Out}^0\mathrm{W}$ with this subgroup. It was observed in part (b) that (S, a_1) generates $\operatorname{Aut}^0(\mathrm{W}, a_1)$. One may easily verify that $\operatorname{Aut}^0(W, a_1) = \operatorname{Out}^0\mathrm{W} \rtimes \langle x_{21} \rangle$, so $\operatorname{Out}^0\mathrm{W}$ has index 2 in $\operatorname{Aut}^0(W, a_1)$, and index $n!$ in $\operatorname{Out}\mathrm{W}$.

Gilman [5] proved that if G is a group and H is a finite-index subgroup of G, then G admits a regular language of normal forms if and only if H does. Thus we have the following.

Corollary 4.3. $\operatorname{Out}\mathrm{W}$ *admits a regular language of normal forms.*

5 Some properties of $\mathcal{L}_I$

In this section we describe some computations involving the rewriting system associated to $\mathcal{L}_I$, the language of normal forms for $\operatorname{Aut}^0\mathrm{W}$ constructed in Section 3, and the conclusions we draw from it.

Lemma 5.1. *Consider the letter* $s := [x_{12}x_{32}x_{13}] \in S$ *and the sequence of normal forms* $\mathsf{U}_k := \big([x_{12}x_{32}][x_{21}x_{31}]\big)^k \in \mathcal{L}_I$*. For each* $k \geq 1$*, the normal form for* $\mathsf{U}_k s$ *is*

$$\mathsf{V}_k := [x_{12}x_{32}x_{13}]\big([x_{13}x_{23}][x_{21}x_{31}][x_{13}x_{23}][x_{12}x_{32}]\big)^k \in \mathcal{L}_I.$$

Fig. 5 depicts U_k *and* V_k *in the Cayley graph* $\Gamma_S(\operatorname{Aut}^0\mathrm{W}_3)$*.*

Example 5.2. Figure 5 illustrates an example of rewriting in $\operatorname{Aut}\mathrm{W}_3$. The figure depicts the rewriting of $\mathsf{U}_k s$ to V_k, with variables as in Computation 5.1. Paths are drawn in the Cayley graph $\Gamma_S(\operatorname{Aut}^0\mathrm{W}_3)$. Each pentagon represents a rewriting step, since each pair consisting of a horizontal edge (from left-to-right) followed a vertical edge is forbidden and the corresponding normal form is spelled by the three edges that complete the pentagon. It is clear that the paths corresponding to U_k and V_k asynchronously fellow-travel with a fellow-traveler constant of 2.

We learn the following from our computation.

Lemma 5.3. *For* $n \geq 3$*, the language* $\mathcal{L}_I$ *is not a language of geodesics and is not an automatic (nor a biautomatic) structure on* $\operatorname{Aut}^0\mathrm{W}$*.*

Proof. The first part of the conclusion is clear because $\mathsf{V}_1 \in \mathcal{L}_I$, V_1 is strictly longer than $\mathsf{U}_1 s$, and both represent the same element of $\operatorname{Aut}^0\mathrm{W}$. To prove the second part of the conclusion, we consider normal forms as paths in the Cayley graph $\Gamma_S(\operatorname{Aut}^0\mathrm{W})$. Note that Property (I2) implies that such paths do not contain loops. If $\mathcal{L}_I \subset S^*$ were part of an automatic structure on G, then the fellow-traveler property and the observation in the previous sentence would imply the existence of a universal upper bound on the difference in the lengths of words $\mathsf{X}, \mathsf{Y} \in \mathcal{L}_I$ for which $\mathsf{X}^{-1}\mathsf{Y} \in S$ (considered as elements of $\operatorname{Aut}^0\mathrm{W}$). But $\mathsf{U}_k, \mathsf{V}_k \in \mathcal{L}_I$, $\mathsf{U}_k^{-1}\mathsf{V}_k \in S$ (considered as elements of $\operatorname{Aut}^0\mathrm{W}$) and $|\mathsf{V}_k| - |\mathsf{U}_k| = 2k + 1$. Thus no such universal upper bound exists and $\mathcal{L}_I$ is not part of an automatic structure on $\operatorname{Aut}^0\mathrm{W}$. □

Remark 5.4. We performed computations which confirmed that, when $n = 3$, the rewriting procedure associated to $\mathcal{L}_I$ is complete, and the language $\mathcal{L}_I$ is part of an asynchronously automatic structure on $\mathrm{Aut}^0\,\mathrm{W}$. To do this, we enumerated all words of the form $\mathsf{U}s$, where $s \in S$ and $U \in \mathcal{L}_I$ and U has length at most 3. We then observed the rewriting process. In each case, after each rewriting step, there was at most one forbidden pair, and the forbidden pair marched unfailingly to the left until there was no forbidden pair left. A forbidden pair was never replaced by the empty word, except perhaps at the first step. These observations are sufficient to draw the conclusions above.

The behavior of the rewriting system is not as simple when $n \geq 4$. We leave unanswered the question of whether or not our languages for these groups are asynchronously automatic structures. As noted in the introduction, it is not known whether or not $\mathrm{Aut}^0\,\mathrm{W}_n$ admits any type of automatic structure in this case.

6 Properties (I2) and (I3)

In this section we complete the proof that $\mathcal{L}_I$ is a Markov language by establishing Properties (I2) and (I3). The proof that Property (I3) holds is the most technical part of the paper.

Recall that, for each element $g \in \mathrm{W}$, there is a unique reduced word $\mathsf{U}_g \in A^*$ which spells g, where a word is reduced if it does not contain a subword of the form $a_i a_i$.

Lemma 6.1. *Property (I2) holds.*

Proof. Let $f \in \mathrm{Aut}^0\,\mathrm{W} \setminus \{\mathrm{id}\}$ and let $s = I(f)$. We must show that $\|sf\| < \|f\|$. For each $i \in \mathbf{n}$ such that $a_i s = a_i$, we have that $a_i sf = a_i f$ and $|a_i sf| = |a_i f|$. Since $f \neq \mathrm{id}$, Property (I1) gives that $s \neq \mathrm{id}$. So there exists some $i \in \mathbf{n}$ such that $a_i s \neq a_i$. Thus it suffices to show that $|a_i sf| < |a_i f|$ for each $i \in \mathbf{n}$ such that $a_i s \neq a_i$.

Let $i \in \mathbf{n}$ be such that $a_i s \neq a_i$. Let $u_i \in \mathrm{W}$ (resp. $w_i \in \mathrm{W}$) be the minimal length element such that $a_i s = u_i a_i \overline{u}_i$ (resp. $a_i f = w_i a_i \overline{w}_i$). Then $a_i(sf) = (u_i f) w_i a_i \overline{w}_i (\overline{u}_i f)$, and $u_i = a_{i_1} a_{i_2} \dots a_{i_k}$ for some $1 \leq k < n$ and some distinct integers $i_1, i_2, \dots i_k \in \mathbf{n} \setminus \{i\}$. The geodesic from v_1 to $F_{a_i s}$ passes through the points

$$F_{a_{i_1}s}, F_{a_{i_2}s}, \dots, F_{a_{i_k}s}$$

in order. By the definition of $I(f)$, the geodesic from v_1 to $F_{a_i f}$ passes through the points

$$F_{a_{i_1}f}, F_{a_{i_2}f}, \dots, F_{a_{i_k}f}$$

in order, and we may subdivide this geodesic into subpaths labeled according to Fig. 6, where the subpaths labeled y_0, y_1, …, y_k are possibly trivial. So we

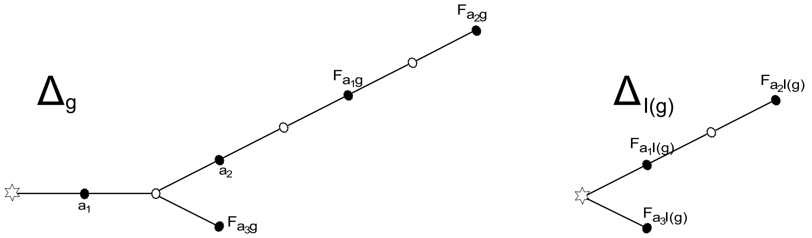

Figure 4: Δ_g and $\Delta_{I(g)}$ for Example 3.9.

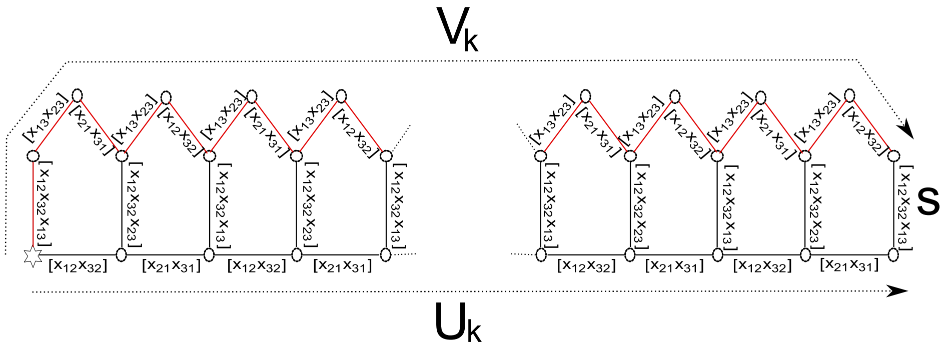

Figure 5: The rewriting for U_k and s as in Lemma 5.1.

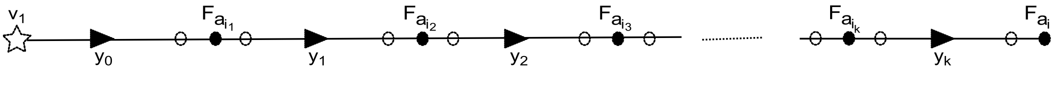

Figure 6: The notation in the proof of Property (I2).

have

$$\begin{array}{rcl}
a_{i_1}f & = & y_0 a_{i_1} \overline{y}_0 \\
a_{i_2}f & = & y_0 a_{i_1} y_1 a_{i_2} \overline{y}_1 a_{i_1} \overline{y}_0 \\
\vdots & \vdots & \vdots \\
a_{i_k}f & = & y_0 a_{i_1} y_1 a_{i_2} \dots y_{k-1} a_{i_k} \overline{y}_{k-1} \dots a_{i_2} \overline{y}_1 a_{i_1} \overline{y}_0
\end{array}$$

and

$$a_i f = y_0 a_{i_1} y_1 a_{i_2} y_2 a_{i_3} \dots y_{k-1} a_{i_k} y_k a_i \overline{y}_k a_{i_k} \overline{y}_{k-1} \dots a_{i_3} \overline{y}_2 a_{i_2} \overline{y}_1 a_{i_1} \overline{y}_0. \qquad (1)$$

In Fig. 7, one can read paths

$$(a_{i_1}f)(a_{i_2}f) \dots (a_{i_k}f)(a_i f)(a_{i_k}f) \dots (a_{i_2}f)(a_{i_1}f)$$

and

$$y_0 y_1 y_2 \cdots y_{k-1} y_k a_i \overline{y}_k \overline{y}_{k-1} \cdots \overline{y}_2 \overline{y}_1 \overline{y}_0$$

with the same end-points. It follows that

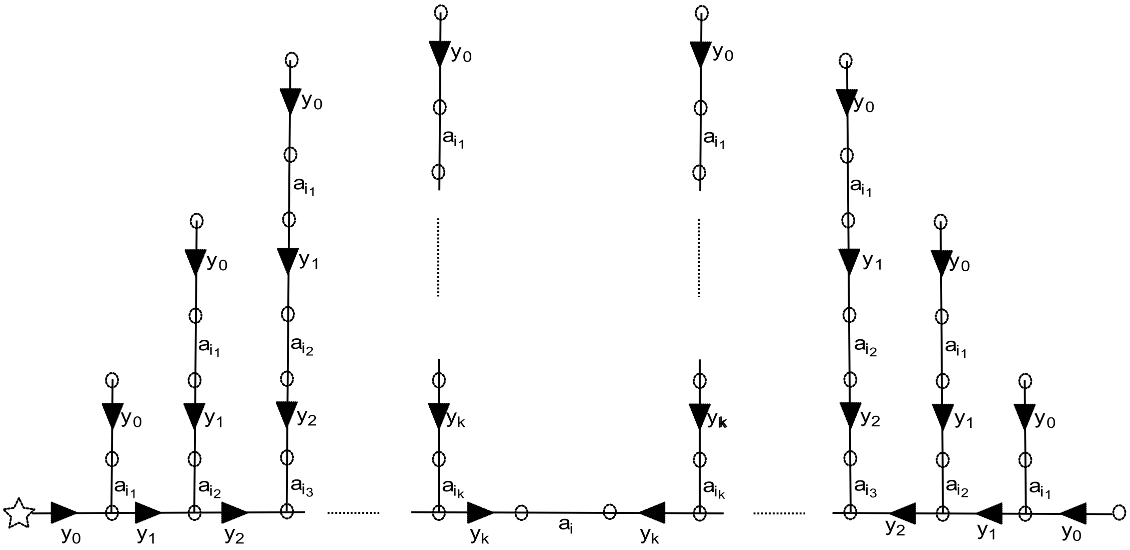

Figure 7: Cancelation in the proof of Property (I2).

$$a_i(sf) = (a_{i_1} a_{i_2} \dots a_{i_k} a_i a_{i_k} \dots a_{i_2} a_{i_1})f = y_0 y_1 y_2 \dots y_{k-1} y_k a_i \overline{y}_k \overline{y}_{k-1} \cdots \overline{y}_2 \overline{y}_1 \overline{y}_0. \qquad (2)$$

Since the right-hand side of Equation (2) (which is possibly unreduced) is obtained from the right-hand side of Equation (1) (which is definitely reduced) by omitting $2k$ letters, we have that $|a_i(sf)| < |a_i f|$. The result follows. $\square$

We now address Property (I3). Our idea of proof is a straight-forward adaptation of that used to prove the analogous statement in [6, Lemma 6.8]. We have been careful to match up the notation as much as possible for the convenience of the reader familiar with the original argument.

Following [6], we develop a graphical way to consider the passage from Δ_s to Δ_{sf}, for $s \in S$ and $f \in \mathrm{Aut}^0 \mathrm{W}$. Recall that the labeled graph Δ_s (resp. Δ_{sf}) is the minimal connected subgraph of Γ which contains v_1 and the fixed points $F_{a_1 s}, \ldots, F_{a_n s}$ (resp. $F_{a_1 sf}, \ldots, F_{a_n sf}$). We construct two more labeled graphs.

1. For each $i \in \mathbf{n}$ we do the following. Let $w_i \in \mathrm{W}$ be the minimal length element such that $a_i f = w_i a_i \overline{w}_i$. There is a unique black vertex in Δ_s with label $F_{a_i s}$, and that black vertex is either the midpoint of an edge or at the end of exactly one half edge. If $F_{a_i s}$ is the midpoint of an edge, then replace this edge by a path with label $w_i a_i \overline{w}_i$ (that is, subdivide the edge into $2\,|w_i| + 1$ edges and label appropriately) and label the black vertex at the midpoint of this path by $F'_{a_i(sf)}$. If $F_{a_i s}$ is the endpoint of exactly one half edge, then replace this half edge by a path with label w_i followed by a half edge, and label the black vertex at the end of the half edge by $F'_{a_i(sf)}$ (that is, subdivide the half edge into $|w_i|$ edges and one half edge and label appropriately). The resulting graph is called Δ'_{sf}.
2. Let Δ''_{sf} be the result of performing Stalling's folding operation [12] on Δ'_{sf}.

These constructions mimic the way in which one might compute the reduced word for $a_i sf$ from the reduced word for each $a_i s$. First one replaces each letter in the reduced word for $a_i s$ by its image under f—compare this to the construction of Δ'_{sf}. Then one performs any necessary cancelation in the result—compare this to the construction of Δ''_{sf}.

Because Δ''_{sf} has been folded, it embeds in Γ. Then Δ_{sf} is isomorphic to the minimal connected subgraph of Δ''_{sf} which contains the image of v_1 and the images of the points $F'_{a_1 sf}, \ldots, F'_{a_n sf}$. Further, $F_{a_i sf} \in \Delta_{sf}$ is the image of $F_{a_i sf} \in \Delta'_{sf}$ for each $i \in \mathbf{n}$.

Example 6.2. In Fig. 8 we illustrate Δ_t, Δ'_{tg}, Δ''_{tg} and Δ_{tg} for $t = x_{21}x_{31} \in S_3$ and $g = x_{21}x_{12}x_{32}x_{23} \in \mathrm{Aut}^0 \mathrm{W}_3$.

For elements $u, u_1, u_2 \in \mathrm{W}$, we write $u \equiv u_1 u_2$ to mean that the reduced word for u is obtained by concatenating the reduced words for u_1 and u_2 (that is, no cancelation is necessary). Let $f \in \mathrm{Aut}^0 \mathrm{W}$ and, for each $i \in \mathbf{n}$, write $w_i \in \mathrm{W}$ for the minimal length element such that $a_i f = w_i a_i \overline{w}_i$. We encode the configuration of the fixed points $F_{a_1 f}, \ldots, F_{a_n f}$ in a function $\epsilon(f)\colon \{(i,k) \mid i, k \in \mathbf{n}, i \neq k\} \to \{0,1\}$ as follows:

> $\epsilon(f)(i,k) = 1$ if and only if the geodesic from v_1 to $F_{a_i f}$ passes through $F_{a_k f}$.

Equivalently,

$\epsilon(f)(i,k) = 1$ if and only if $w_i \equiv w_k a_k w_i'$ for some (possibly trivial) element w_i'.

Following [6], we write $\epsilon_{ik}(f) := \epsilon(f)(i,k)$.

Example 6.3. For $g = x_{21}x_{12}x_{32}x_{23} \in \mathrm{Aut}^0 \mathrm{W}_3$ (with Δ_g as depicted in Fig. 4), we have $\epsilon_{21}(g) = 1$, and $\epsilon_{ik}(g) = 0$ for all other pairs of distinct integers $i, k \in \{1, 2, 3\}$.

It is immediate that the following properties hold for each $s, s' \in S$ and each $f \in \mathrm{Aut}^0 \mathrm{W} \setminus \{\mathrm{id}\}$:

1. $s = s'$ if and only if $\epsilon(s) = \epsilon(s')$;
2. $\epsilon(f) = \epsilon(I(f))$ (that is, ϵ depends only on the initial letter of f).

It follows that $I(sf) = s$ if and only if $\epsilon(sf) = \epsilon(s)$. Further, it follows that Property (I3) is proved if we can show that the functions $\epsilon(s)$ and $\epsilon(f)$ determine whether or not the equality $\epsilon(sf) = \epsilon(s)$ holds. Lemma 6.4 below achieves this, and thus completes the proof that $\mathcal{L}_I$ is a Markov language.

To state the criteria on $\epsilon(s)$ and $\epsilon(f)$ under which the equality $\epsilon(sf) = \epsilon(s)$ holds, it is convenient to introduce two more terms to our vocabulary and some notation. We say that $k \in \mathbf{n}$ is (s,f)-*active*, or just *active*, if $\epsilon_{ik}(f) = 1$ for each $i \in \mathbf{n}$ such that $d_\Gamma(F_{a_i s}, F_{a_k s}) = 1$; that is, $w_k a_k$ is an initial subword of w_i whenever the edge labeled $F_{a_i s}$ is adjacent to the edge labeled $F_{a_k s}$. We say that $k \in \mathbf{n}$ is (s,f)-*passive*, or just *passive*, if $\epsilon_{ik}(f) = 0$ for each $i \in \mathbf{n}$ such that $d_\Gamma(F_{a_i s}, F_{a_k s}) = 1$; that is, $w_k a_k$ is not an initial subword of w_i whenever the edge labeled $F_{a_i s}$ is adjacent to the edge labeled $F_{a_k s}$.

Lemma 6.4. *Let $s \in S$ and $f \in \mathrm{Aut}^0 \mathrm{W} \setminus \{id\}$. Then $\epsilon(sf) = \epsilon(s)$ if and only if the following properties both hold for each $k \in \mathbf{n}$:*

(a) either k is (s,f)-active or k is (s,f)-passive;

(b) if $\epsilon_{ki}(s) = 0$ for each $i \in \mathbf{n} \setminus \{k\}$ (equivalently, $a_k s = a_k$), then k is (s,f)-passive.

Proof. As above, we shall write $w_i \in \mathrm{W}$ for the minimal length element such that $a_i f = w_i a_i \overline{w}_i$.

Assume that Properties (a) and (b) hold. We must show that $\epsilon(sf) = \epsilon(f)$; that is, Δ_{sf} and Δ_s have the same configuration of fixed points. We thank the anonymous referee for the following intuitive description of our argument: when passing from Δ_s to Δ_{sf}, Property (a) ensures that the configuration of fixed pints around the edge containing $F_{a_k s}$ is “flipped” (the fixed points on each side fold to be on the other side) if k is active, and is unchanged if k is passive; Property (b) ensures that the basepoint is not flipped across a fixed point. Thus Property (a) ensures that the tree and the labels $F_{a_k s}$ remain unchanged, while Property (b) ensures that the basepoint remains in the same position on the tree.

It is immediate from the definitions that if $\epsilon_{ik}(f) = 1$, then $\epsilon_{ki}(f) = 0$. It follows that if $i, k \in \mathbf{n}$ are both active, then $d(F_{a_i s}, F_{a_k s}) \neq 1$.

Assume that $k \in \mathbf{n}$ is active. In Fig. 9 we follow the passage from Δ_s to Δ''_{sf} for a neighborhood of $F_{a_k s}$. The fixed points which are unit distance from $F_{a_k s}$ are labeled by $F_{a_{i_1} s}, \ldots, F_{a_{i_p} s}, F_{a_{j_1} s}, \ldots, F_{a_{j_q} s}$. Since k is active, we have from the paragraph above that $i_1, \ldots, i_p, j_1, \ldots, j_q$ are passive. It follows that $w_{i_1} \equiv w_k a_k w'_{i_1}$ for some (possibly trvial) element w'_{i_1}; we define elements $w'_{i_2}, \ldots, w'_{i_p}, w'_{j_1}, \ldots, w'_{j_q}$ analogously.

In Diagram (A) we show a neighborhood of $F_{a_k s}$ in Δ_s.

In Diagram (B) we show a neighborhood of $F'_{a_k sf}$ in Δ'_{sf}. To pass from (A) to (B) we have replaced each edge labeled a_ℓ by a path $w_\ell a_\ell \overline{w}_\ell$. It is convenient to draw a directed edge with label w_ℓ rather than depicting the undirected edges with labels that spell w_ℓ.

Diagrams (B) and (C) show the same neighborhood of Δ'_{sf}. The difference is that in Diagram (C) we have subdivided the paths $w_{i_1}, \ldots, w_{i_p}, w_{j_1}, \ldots, w_{j_q}$ into the subpaths $w_k a_k w'_{i_0}, \ldots, w_k a_k w'_{i_p}, w_k a_k w'_{j_1}, \ldots, w_k a_k w'_{j_q}$ respectively.

The passage from Diagram (C) to Diagram (D) is achieved by some folding. We note that the configuration of fixed points in Diagram (D) matches the configuration of fixed points in Diagram (A). Now, Diagram (D) may not be an immersion (that is, we may not be finished folding) since some folding of the paths w'_{i_x} and w'_{i_y} may be required before the picture will embed in Δ''_{sf}. However, such folding will not change the configuration of fixed points because, since i_x is passive, $\epsilon_{i_y i_x}(f) = 0$ and $w'_{i_x} a_{i_x}$ is not an initial subword of w'_{i_y}. Similarly, some of the paths w'_{j_x} and w'_{j_y} may require folding, but such folding will not change the configuration of fixed points. We shall write (D') for the final result of folding (D).

In Fig. 10 we follow the passage from Δ_s to Δ''_{sf} for a neighborhood of a white vertex for which the incident edges contain fixed points $F_{a_{i_1} s}, \ldots, F_{a_{i_p} s}$ and $i_1, \ldots, i_p$ are passive. In Diagram (E) we show the neighborhood in Δ_s. In Diagram (F) we show the corresponding neighborhood in Δ'_{sf}. To pass from Diagram (E) to Diagram (F) we have replaced each edge labeled a_ℓ by a path $w_\ell a_\ell \overline{w}_\ell$. Some of the paths w_{i_x} and w_{i_y} may require folding before the picture will embed in Δ''_{sf}, but, as in the paragraph above, such folding will not change the configuration of fixed points. We write (F') for the final result of folding (F).

One may construct Δ_{sf} by 'piecing together' neighborhoods like (D') and (F'). Piecing together is achieved by identifying fixed points with the same label $F_{a_i sf}$, as appropriate (note that these are always at the extremities of the neighborhoods). It follows that the configurations of fixed points in Δ_{sf} and Δ_s match. Condition (b) ensures that the position of the basepoint remains unchanged. Hence $\epsilon(sf) = \epsilon(f)$ as required.

Now suppose that Property (b) fails. Let $i, k \in \mathbf{n}$ be distinct integers such that $a_k s = a_k$, $d_\Gamma(F_{a_k s}, F_{a_i s}) = 1$ and $\epsilon_{ik}(f) = 1$. It follows that $w_i \equiv w_k a_k w'_i$ for some (possibly trivial) element w'_i. In particular, we note if w'_i is nontrivial, then the first letter is not a_k. Further, we have that $a_k sf = a_k f = w_k a_k \overline{w}_k$ and

$d_\Gamma(F_{a_i s}, v_1) \in \{1/2, 3/2\}$. If $d_\Gamma(F_{a_i s}, v_1) = 1/2$, then

$$a_i sf = a_i f \equiv w_i a_i \overline{w}_i \equiv w_k a_k w_i' a_i \overline{w}_i' a_k \overline{w}_k;$$

it follows that the geodesic from v_1 to $F_{a_i sf}$ passes through $F_{a_k sf}$ and $\epsilon_{ik}(sf) = 1 \neq 0 = \epsilon_{ik}(s)$. If $d_\Gamma(F_{a_i s}, v_1) = 3/2$, then

$$a_i sf = (a_k a_i a_k)f = w_k a_k \overline{w}_k w_k a_k w_i' a_i \overline{w}_i' a_k \overline{w}_k w_k a_k \overline{w}_k = w_k w_i' a_i \overline{w}_i' \overline{w}_k;$$

it follows that the geodesic from v_1 to $F_{a_i sf}$ does not pass through $F_{a_k sf}$ and $\epsilon_{ik}(sf) = 0 \neq 1 = \epsilon_{ik}(s)$.

Now suppose that Property (a) fails. Let $k \in \mathbf{n}$ be an integer which is neither active nor passive. So there exist distinct integers $i, j \in \mathbf{n}$, with $d_\Gamma(F_{a_k s}, F_{a_i s}) = d_\Gamma(F_{a_k s}, F_{a_j s}) = 1$ and such that $\epsilon_{ik}(f) = 1 \neq 0 = \epsilon_{jk}(f)$. It follows immediately that one of the following holds:

(X) $d_\Gamma(F_{a_i s}, F_{a_j s}) = 1$;

(Y) $d_\Gamma(F_{a_i s}, F_{a_j s}) = 2$.

It is immediate from the definitions that, for distinct integers $a, b, c \in \mathbf{n}$: if $\epsilon_{ba}(f) = 1$, then $\epsilon_{ab}(f) = 0$; if $\epsilon_{ba}(f) = \epsilon_{cb}(f) = 1$, then $\epsilon_{ca}(f) = 1$; and if $\epsilon_{ca}(f) = \epsilon_{cb}(f) = 1$, then $\epsilon_{ba}(f) = 1$ or $\epsilon_{ab}(f) = 1$. Our hypotheses imply that one of the following holds:

(i) $\epsilon_{kj}(f) = \epsilon_{ij}(f) = \epsilon_{ji}(f) = 0$;

(ii) $\epsilon_{kj}(f) = \epsilon_{ij}(f) = 1$ and $\epsilon_{ji}(f) = 0$.

The four cases (combinations of (X) or (Y) with (i) or (ii)) may be considered separately. In each case, to show that $\epsilon(sf) \neq \epsilon(s)$ it is enough to show that something about the configuration of fixed points $\{F_{a_i sf}, F_{a_j sf}, F_{a_k sf}\}$ does not match the configuration of the fixed points $\{F_{a_i s}, F_{a_j s}, F_{a_k s}\}$. For example, it would suffice to show that the geodesic from $F_{a_i sf}$ to $F_{a_j sf}$ passes through $F_{a_k sf}$, but the geodesic from $F_{a_i s}$ to $F_{a_j s}$ does not pass through $F_{a_k s}$.

Suppose that (X) and (i) hold. The following argument is illustrated in Fig. 11. Since $\epsilon_{ik}(f) = 1$, $w_i \equiv w_k a_k w_i'$ for some (possibly trivial) element w_i'. Since (X) holds, the geodesic from $F_{a_i s}$ to $F_{a_j s}$ does not pass through $F_{a_k s}$. In Diagram (H) we show a connected subgraph of Δ_s. In Diagram (I) we show a connected subgraph in Δ'_{sf}. To pass from (H) to (I) we have replaced each half edge labeled $F_{a_\ell s}$ by a path w_ℓ and a half edge labeled $F_{a_\ell s}$. Diagrams (I) and (J) show the same neighborhood of Δ'_{sf}. The difference is that in Diagram (J) we have subdivided the path w_i into the subpaths $w_k a_k w_i'$. The passage from Diagram (J) to Diagram (K) is achieved by some folding. We note that the configuration of fixed points in Diagram (K) does not match the configuration of fixed points in Diagram (H), as the geodesic from $F''_{a_i sf}$ to $F''_{a_j sf}$ passes through $F''_{a_k sf}$. Now, Diagram (K) may not be an immersion (that is, we may not be finished folding) since some folding of the paths labeled w_j and w_k may be required before the picture will embed in Δ''_{sf}. However, such folding will not change the configuration of fixed points because, since $\epsilon_{jk}(f) = \epsilon_{kj}(f) = 0$, $w_j a_j$ is not an initial subword of w_k and $w_k a_k$ is not an initial subword of w_j.

We include figures (Figures 12, 13 and 14) which illustrate the folding in the remaining three cases, but omit the commentary. □

Acknowledgments

We thank the anonymous referee for suggesting a number of improvements to the exposition of this paper. We dedicate this paper to Mauricio Gutierrez, our colleague and friend.

References

[1] R. Charney. Geodesic automation and growth functions for Artin groups of finite type. *Math. Ann.*, 301(2):307–324, 1995.

[2] D. J. Collins. Palindromic automorphisms of free groups. In *Combinatorial and geometric group theory (Edinburgh, 1993)*, volume 204 of *London Math. Soc. Lecture Note Ser.*, pages 63–72. Cambridge Univ. Press, Cambridge, 1995.

[3] D. B. A. Epstein, J. W. Cannon, D. F. Holt, S. V. F. Levy, M. S. Paterson, and W. P. Thurston. *Word processing in groups.* Jones and Bartlett Publishers, Boston, MA, 1992.

[4] N. D. Gilbert. Presentations of the automorphism group of a free product. *Proc. London Math. Soc. (3)*, 54(1):115–140, 1987.

[5] R. H. Gilman. Groups with a rational cross-section. In *Combinatorial group theory and topology (Alta, Utah, 1984)*, volume 111 of *Ann. of Math. Stud.*, pages 175–183. Princeton Univ. Press, Princeton, NJ, 1987.

[6] M. Gutiérrez and S. Krstić. Normal forms for basis-conjugating automorphisms of a free group. *Internat. J. Algebra Comput.*, 8(6):631–669, 1998.

[7] M. Gutierrez, A. Piggott, and K. Ruane. On the automorphisms of a graph product of abelian groups. To appear in Groups Geom. Dyn., August 2007.

[8] C. Jensen, J. McCammond, and J. Meier. The Euler characteristic of the Whitehead automorphism group of a free product. *Trans. Amer. Math. Soc.*, 359(6):2577–2595 (electronic), 2007.

[9] D. McCullough and A. Miller. Symmetric automorphisms of free products. *Mem. Amer. Math. Soc.*, 122(582):viii+97, 1996.

[10] A. Piggott, K. Ruane, and G. S. Walsh. The automorphism group of the free group of rank two is a CAT(0) group. To appear in Mich. Math. J., November 2008.

[11] J. S. Rose. Automorphism groups of groups with trivial centre. *Proc. London Math. Soc. (3)*, 31(2):167–193, 1975.

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

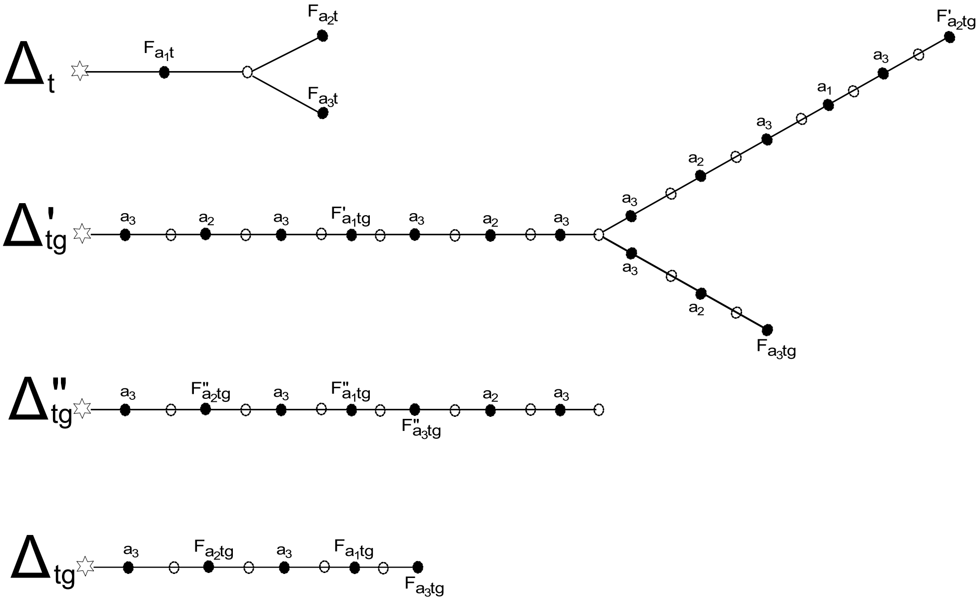

Figure 8: The passage from Δ_t to Δ_{tg} corresponding to Example 6.2.

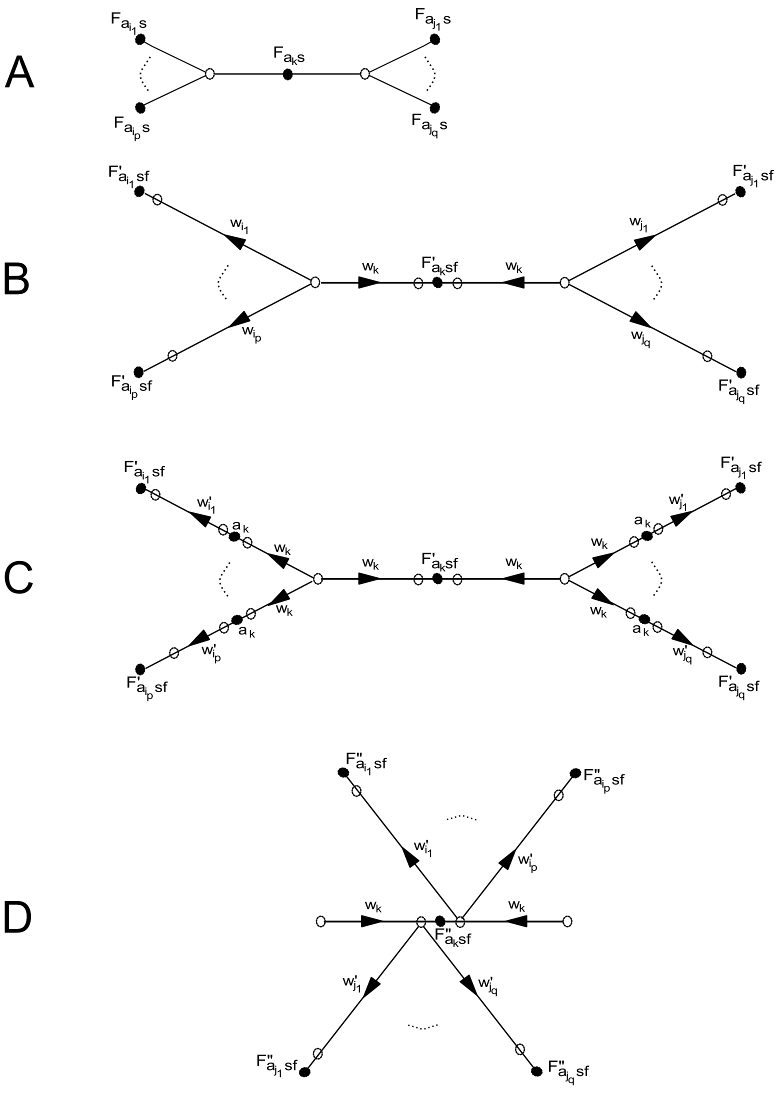

Figure 9: An illustration of folding when k is active.

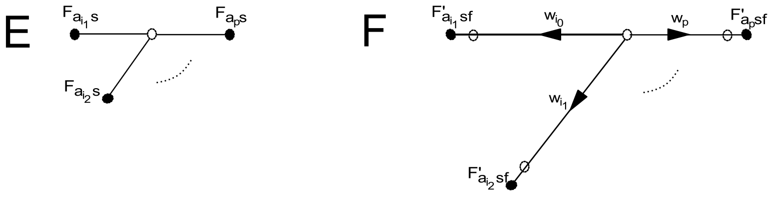

Figure 10: An illustration of folding when $i_1, i_2, \dots i_p$ are passive.

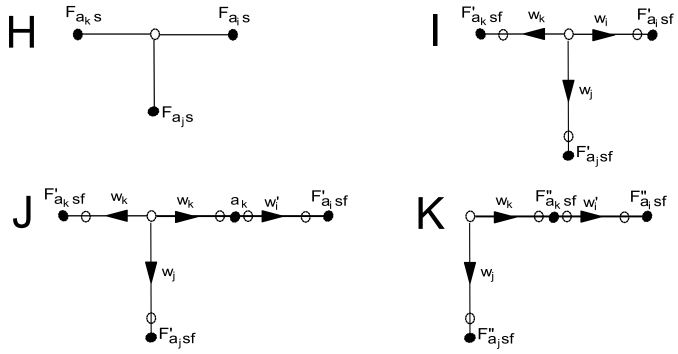

Figure 11: An illustration of folding in the case that (X) and (i) hold.

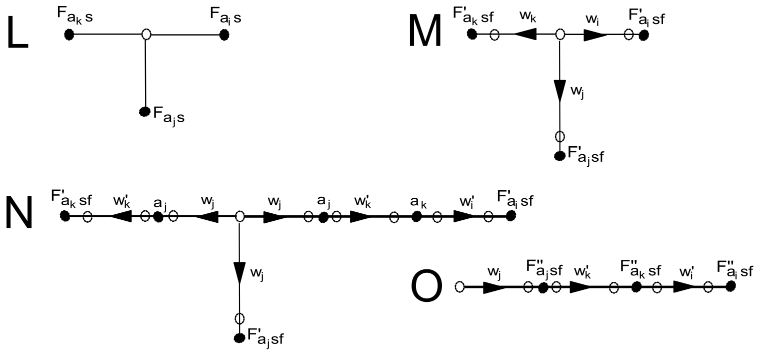

Figure 12: An illustration of folding in the case that (X) and (ii) hold.

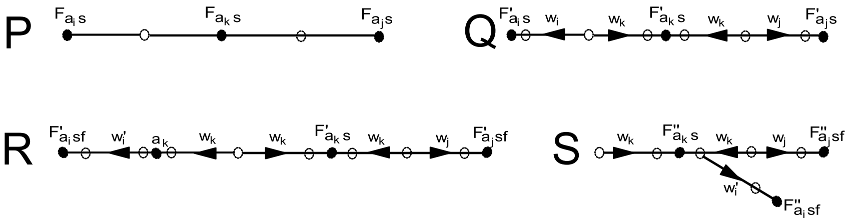

Figure 13: An illustration of folding in the case that (Y) and (i) hold.

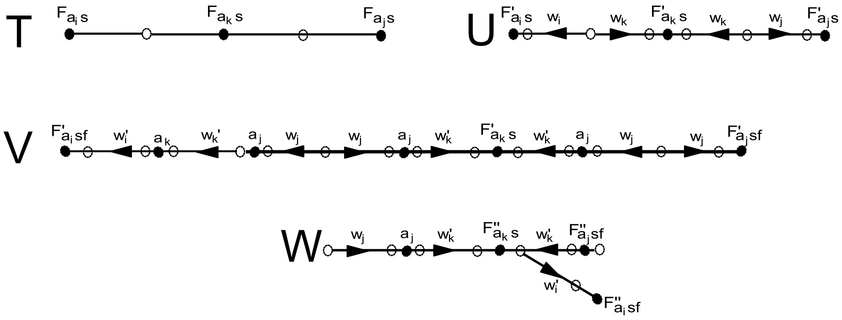

Figure 14: An illustration of folding in the case that (Y) and (ii) hold.