A SIMPLER PROOF OF MCCOOL'S THEOREM

SINA ABBASI

Abstract. In 1947, Artin found a representation of the braid group on n strands in the automorphism group of the free group with n generators, and proved his representation is faithful. In 1986, McCool proved a theorem giving a presentation for the group of basis-conjugating automorphisms of the free group with n generators. McCool's theorem can be reinterpreted as the statement that a natural analogue of Artin's representation of the braid group is a faithful representation of the pure welded braid group on n strands. However, the proof of McCool's theorem uses Whitehead's peak reduction lemma, which has no interpretation in the context of welded braids. In this paper, we apply a version of Whitehead's peak reduction lemma suited for welded braids in order to obtain a simpler proof of McCool's theorem.

1. Background

Let B_n denote the braid group on n strands, PWB_n denote the pure welded braid group on n strands, $\mathcal{F}_n$ denote the free group on n generators, and $\text{Aut}(\mathcal{F}_n)$ denote the automorphism group of $\mathcal{F}_n$.

In 1947, Artin [A] found a representation of B_n in $\text{Aut}(\mathcal{F}_n)$ and proved the representation is faithful. In 1986, McCool [M2] proved a theorem giving a presentation for the group of basis-conjugating automorphisms in $\text{Aut}(\mathcal{F}_n)$. The statements of these results are actually closely related, since McCool's 1986 theorem can be interpreted as saying that a natural analogue of Artin's representation of B_n is a faithful representation of PWB_n. However, the proofs of these results appear to be largely independent of each other.

Indeed, the ideas in [M2] trace back to Whitehead's 1936 paper [W], which settled the problem of algorithmically determining whether or not there exists an automorphism ϕ of $\mathcal{F}_n$ with $\phi(u) = u'$, given two words $u, u' \in \mathcal{F}_n$. Whitehead's proof is topological, relating the problem to three-dimensional CW complexes.

In 1958, Rapaport [R] was able to eliminate the use of topology in Whitehead's proof: she instead worked directly with elements and automorphisms of $\mathcal{F}_n$. In doing so, Rapaport showed how the crux of Whitehead's proof is actually the proof of a statement, now called "Whitehead's peak reduction lemma", about the finite generating set of $\text{Aut}(\mathcal{F}_n)$, now called the set of "Whitehead automorphisms", that Whitehead introduced in [W]. In 1974, Higgins and Lyndon [HL] simplified

Date: September 1, 2020.

Rapaport's proof of the peak reduction lemma considerably, while still working directly with elements and automorphisms of $\mathcal{F}_n$.

Shortly after, in 1975, McCool [M1] carefully analyzed Higgins and Lyndon's proof of Whitehead's peak reduction lemma, and used its content to prove a theorem giving a presentation for the stabilizer in $\mathrm{Aut}(\mathcal{F}_n)$ of a tuple of cyclic words. Finally, in 1986, McCool used his theorem from [M1] to prove the aforementioned theorem in [M2], recognizing that the group of basis-conjugating automorphisms in $\mathrm{Aut}(\mathcal{F}_n)$ is the stabilizer of a certain tuple of cyclic words.

Now, as mentioned earlier, McCool's theorem in [M2] can be interpreted as saying that the natural analogue of Artin's representation of B_n is a faithful representation of PWB_n. However, the issue is that, in the context associated with PWB_n, McCool's proof no longer has any direct meaning. Indeed, McCool's proof in [M2] relies on his theorem in [M1] in an essential way, and his theorem in [M1] is a general result with no interpretation in the context of welded braids. Moreover, McCool proves his theorem in [M1] using Whitehead's peak reduction lemma, which involves Whitehead automorphisms, and Whitehead automorphisms have no interpretation in the context of welded braids either.

What led to the creation of this paper was reading Bar-Natan, Dancso and van der Veen's paper [BNDV], where it is proved that a natural analogue of Artin's representation of B_n is a complete invariant for virtual braids. The method of proof there is an application of the "diamond lemma", and the proof stays completely within the context of virtual braids. The similarities between the diamond lemma and Whitehead's peak reduction lemma suggested to us that it should be possible to prove McCool's 1986 theorem by using a version of Whitehead's peak reduction lemma suited to the context of PWB_n. This did turn out to be possible, and the purpose of this paper is to show how it leads to a simpler proof of the theorem in [M2].

While in the process of writing this paper, we discovered that a proof with the same ideas as ours, though presented differently, exists and is the main subject of [FRR]. Interestingly, [FRR] does not cite any of the papers mentioned above, so it appears that the story told above is only one possible path to our proof of McCool's theorem from [M2].

2. Introduction

2.1. **Notation for words.** Suppose S is a given set of symbols. We let S^{-1} denote the set consisting of the same symbols as S, but with the superscript $^{-1}$ added to each. We call a pair of symbols consisting of a symbol s in S along with its corresponding symbol s^{-1} in S^{-1} a *pair of inverses.* In addition, we naturally call each symbol in a pair of inverses the *inverse* of the other symbol.

We let $\mathcal{W}(S)$ denote the set of words with letters consisting of symbols from $S \cup S^{-1}$. We call a word in $\mathcal{W}(S)$ *reduced* if no two adjacent letters in the word constitute a pair of inverses. Given a word $w \in \mathcal{W}(S)$, we can repeatedly survey w

to check if it contains two adjacent letters constituting a pair of inverses, deleting the pair if it does. Such a process will always end with a reduced word after finitely many deletions. We call such a process *reducing* w. Note that there may be several ways of reducing w, but it is straightforward to verify that they all result in the same reduced word.

We define an equivalence relation $\sim$ on $\mathcal{W}(S)$, where for $w_1, w_2 \in \mathcal{W}(S)$, $w_1 \sim w_2$ if reducing w_1 and reducing w_2 result in the same reduced word. We let $\mathcal{F}(S)$ denote the set of equivalence classes of $\mathcal{W}(S)$ under this equivalence relation.

Given two words $w_1, w_2 \in \mathcal{W}(S)$, we let $w_1 w_2$ denote the result of appending w_2 to the right end of w_1, forming a new word in $\mathcal{W}(S)$. The map $(w_1, w_2) \mapsto w_1 w_2$ is an associative, non-commutative, binary operation on $\mathcal{W}(S)$. Also, given a word $w \in \mathcal{W}(S)$, we let $w^{-1} \in \mathcal{W}(S)$ denote the word obtained from w after reversing the order of its letters, and then replacing each letter with its inverse.

It is straightforward to check that the binary operation $(w_1, w_2) \mapsto w_1 w_2$ respects equivalence classes. Therefore, it translates into a well-defined binary operation on $\mathcal{F}(S)$. It is well-known that $\mathcal{F}(S)$ together with this binary operation form a group, typically called the free group generated by the symbols in S. Furthermore, the operation $w \mapsto w^{-1}$ on $\mathcal{W}(S)$ defined above respects equivalence classes as well, and the operation on $\mathcal{F}(S)$ it translates to is exactly the inverse operation coming from the aforementioned group structure on $\mathcal{F}(S)$.

In this paper, We let $\mathcal{W}_n$ denote $\mathcal{W}(\{x_1, \ldots, x_n\})$, $\mathcal{F}_n$ denote $\mathcal{F}(\{x_1, \ldots, x_n\})$, and $\mathrm{Aut}(\mathcal{F}_n)$ denote the automorphism group of $\mathcal{F}_n$. Moreover, throughout this paper, all words should be understood as elements of $\mathcal{W}(S)$ by default, and therefore any operations on words in an expression should be understood as the aforementioned operations on $\mathcal{W}(S)$. The only caveat is that we will not introduce any additional notation to separate a word in $\mathcal{W}(S)$ from its equivalence class in $\mathcal{F}(S)$, but it will be clear from context when the latter should be considered instead of the former.

2.2. The pure welded braid group and automorphisms of $\mathcal{F}_n$.

Let Σ_n denote the set of symbols $\{\sigma_{ij} : 1 \leq i, j \leq n, i \neq j\}$. The pure welded braid group on n strands, denoted PWB_n, is the group $\mathcal{F}(\Sigma_n)$, modulo the relations:

(A) $\sigma_{ij}\sigma_{k\ell} = \sigma_{k\ell}\sigma_{ij}$ for i, j, k, ℓ distinct.

(B) $\sigma_{ij}\sigma_{ik} = \sigma_{ik}\sigma_{ij}$ for i, j, k distinct.

(C) $\sigma_{ij}\sigma_{ik}\sigma_{jk} = \sigma_{jk}\sigma_{ik}\sigma_{ij}$ for i, j, k distinct.

The pure welded braid group is most often understood in a diagrammatic context; however, the definition above will suffice for our purposes. For readers familiar with the diagrammatic context, we note that relation (A) corresponds to the "locality" property of welded braid diagrams, relation (B) corresponds to the fact that overcrossings commute, and relation (C) corresponds to type-III Reidemeister moves.

Let $\phi : \Sigma_n \cup \Sigma_n^{-1} \to \mathrm{Aut}(\mathcal{F}_n)$ denote the map with $\phi(\sigma_{ij})$ acting on the standard basis of $\mathcal{F}_n$ following the rule:

$$\phi(\sigma_{ij})(x_k) = \begin{cases} x_k & \text{if } k \neq j \\ x_i x_j x_i^{-1} & \text{if } k = j, \end{cases}$$

and with $\phi(\sigma_{ij}^{-1})$ acting on the standard basis of $\mathcal{F}_n$ following the rule:

$$\phi(\sigma_{ij}^{-1})(x_k) = \begin{cases} x_k & \text{if } k \neq j \\ x_i^{-1} x_j x_i & \text{if } k = j. \end{cases}$$

Note that $\phi(\sigma_{ij})$ and $\phi(\sigma_{ij}^{-1})$ are certainly homomorphisms from $\mathcal{F}_n$ to $\mathcal{F}_n$. Since they are additionally inverses of each other, they are indeed automorphisms of $\mathcal{F}_n$.

The relations (A), (B) and (C) are preserved under ϕ. Therefore, ϕ extends to a group homomorphism $\Phi : PWB_n \to \mathrm{Aut}(\mathcal{F}_n)$ which is equal to ϕ on the generating set $\Sigma_n \cup \Sigma_n^{-1}$ of PWB_n.

Hereafter, we will suppress the symbol ϕ, instead having $\sigma_{ij}^{\pm 1}$ stand both for a formal symbol in $\Sigma_n \cup \Sigma_n^{-1}$ and for the automorphism $\phi(\sigma_{ij}^{\pm 1})$ of $\mathcal{F}_n$ defined above.

2.3. **Basis-conjugating automorphisms.**

The automorphism $\sigma_{ij}^{\pm 1}$ has the property that it maps each generator x_i of $\mathcal{F}_n$ to a conjugate wx_iw^{-1} of x_i in $\mathcal{F}_n$. We will call automorphisms with this property *basis-conjugating*, and we denote the set of all basis-conjugating automorphisms of $\mathcal{F}_n$ by BCA_n.

The property of being basis-conjugating is preserved under composition of automorphisms. Therefore, since Φ maps $\Sigma_n \cup \Sigma_n^{-1}$ into BCA_n, the image of Φ is contained in BCA_n.

3. The proof of McCool's theorem

The theorem McCool proved in [M2] is equivalent to the following theorem.

Theorem (McCool). *The map $\Phi : PWB_n \to Aut(\mathcal{F}_n)$ defined earlier is injective, and its image is BCA_n.*

The proof of this theorem will be divided into multiple sections. In the section 3.1, we simply introduce some notation; in section 3.2, we prove a lemma fundamental to the overall proof of the theorem; in section 3.3, we prove that the image of Φ is BCA_n; and in section 3.4, we show that Φ is injective.

3.1. **Complexity of words and automorphisms.**

Given a word w in $\mathcal{W}_n$, we refer to the number of letters in w as its *length*, and we let $\ell(w)$ denote the length of the unique reduced word that results from reducing w.

For a basis-conjugating automorphism τ, we let $c(\tau) = \sum_{i=1}^{n} \ell(\tau(x_i)) - n$. Intuitively, we think of $c(\tau)$ as the *complexity* of the automorphism τ.

3.2. Effect of precomposing with a generator on complexity.

Lemma 1. *Let τ be a basis-conjugating automorphism, and for $1 \leq k \leq n$, let $w_k \in \mathcal{W}_n$ be such that $w_k x_k w_k^{-1}$ is reduced and $\tau(x_k) = w_k x_k w_k^{-1}$. Then, if w_j starts with $w_i x_i^{\mp 1}$, $c(\tau \circ \sigma_{ij}^{\pm 1}) < c(\tau)$. Otherwise, $c(\tau \circ \sigma_{ij}^{\pm 1}) > c(\tau)$.*

Proof. Note that $\tau \circ \sigma_{ij}^{\pm 1}$ acts on the standard basis of $\mathcal{F}_n$ following the rule:

$$(\tau \circ \sigma_{ij}^{\pm 1})(x_k) = \begin{cases} w_k x_k w_k^{-1} & \text{if } k \neq j \\ w_i x_i^{\pm 1} w_i^{-1} w_j x_j w_j^{-1} w_i x_i^{\mp 1} w_i^{-1} & \text{if } k = j \end{cases}.$$

Hence,

$$c(\tau \circ \sigma_{ij}^{\pm 1}) = c(\tau) + \ell(w_i x_i^{\pm 1} w_i^{-1} w_j x_j w_j^{-1} w_i x_i^{\mp 1} w_i^{-1}) - \ell(w_j x_j w_j^{-1}) \tag{1}$$

Let $W = w_i x_i^{\pm 1} w_i^{-1} w_j x_j w_j^{-1} w_i x_i^{\mp 1} w_i^{-1}$. We will relate $\ell(W)$ to $\ell(w_j x_j w_j^{-1})$ so that we can apply (1). Let $u \in \mathcal{W}_n$ and $m \in \mathbb{Z}$ be such that $u x_j$ and $u x_j^{-1}$ are reduced (i.e. u is reduced and does not end in x_j or x_j^{-1}) and $w_i x_i^{\pm 1} w_i^{-1} w_j = u x_j^m$. For the purpose of computing $\ell(W)$, we consider the way of reducing W which consists of the following steps:

(i) The deletion of all pairs of adjacent inverses in the instance of $w_i x_i^{\pm 1} w_i^{-1} w_j$ in the expression defining W, resulting in the word $u x_j^m x_j w_j^{-1} w_i x_i^{\mp 1} w_i^{-1}$.

(ii) The deletion of all pairs of adjacent inverses in the instance of $w_j^{-1} w_i x_i^{\mp 1} w_i^{-1}$ in the expression for the word obtained in the previous step, resulting in the word $u x_j^m x_j x_j^{-m} u^{-1}$.

(iii) The deletion of all pairs of adjacent inverses in the instance of $x_j^m x_j x_j^{-m}$ in the expression for the word obtained in the previous step, resulting in the reduced word $u x_j u^{-1}$.

There are now two cases to consider.

Case 1: w_j starts with $w_i x_i^{\mp 1}$.

Then, in step (i) above, at least $\ell(w_i) + 1$ deletions are performed, since each letter in the instance of $x_i^{\pm 1} w_i^{-1}$ in $w_i x_i^{\pm 1} w_i^{-1} w_j$ can be deleted as part of a pair of adjacent inverses, with the other letter in the pair coming from the instance of w_j in $w_i x_i^{\pm 1} w_i^{-1} w_j$. By symmetry, at least $\ell(w_i) + 1$ deletions are performed in step (ii) as well.

Hence, there are at least $4\ell(w_i) + 4$ letters deleted in our above way of reducing W. Since the length of W is $4\ell(w_i) + 2 + \ell(w_j x_j w_j^{-1})$ (note we chose w_j so that $w_j x_j w_j^{-1}$ is reduced), it follows that $\ell(W) < \ell(w_j x_j w_j^{-1})$. Therefore, by (1), $c(\tau \circ \sigma_{ij}^{\pm 1}) < c(\tau)$.

Case 2: w_j does not start with $w_i x_i^{\mp 1}$.

By our choices of w_i and w_j, $w_i x_i^{\pm 1} w_i^{-1}$ and $w_j x_j w_j^{-1}$ are both reduced. It follows that each deletion in step (i) must involve one letter from the instance of $w_i x_i^{\pm 1} w_i^{-1}$ in $w_i x_i^{\pm 1} w_i^{-1} w_j$ and one letter from the instance of w_j. Since we are assuming w_j does not start with $w_i x_i^{\mp 1}$, it follows that each deletion in step (i) involves a letter from the instance of w_i^{-1} in $w_i x_i^{\pm 1} w_i^{-1} w_j$. By symmetry, each deletion in step (ii) involves a letter from the instance of w_i in $w_j^{-1} w_i x_i^{\mp 1} w_i^{-1}$.

Since $w_j x_j w_j^{-1}$ is reduced, it follows that the letters in the instances of x_j^m and x_j^{-m} in $x_j^m x_j x_j^{-m}$ from step (iii) are from the instances of w_i^{-1} in $w_i x_i^{\pm 1} w_i^{-1} w_j$ and w_i in $w_j^{-1} w_i x_i^{\mp 1} x_i^{-1}$ respectively. Therefore, we have shown that each deletion throughout steps (i) to (iii) involves a letter from either the instance of w_i^{-1} following $x_i^{\pm 1}$ in the expression defining W or the instance of w_i preceding $x_i^{\mp 1}$.

Overall, it follows that there are at most $2\ell(w_i)$ deletions of pairs of adjacent inverses in our above way of reducing W, and so at most $4\ell(w_i)$ letters deleted. Since the length of W is $4\ell(w_i) + 2 + \ell(w_j x_j w_j^{-1})$, it follows that $\ell(W) > \ell(w_j x_j w_j^{-1})$. Therefore, by (1), $c(\tau \circ \sigma_{ij}^{\pm 1}) > c(\tau)$.

The cases above and their conclusions form a complete proof of the lemma. □

3.3. **The image of Φ.** Suppose for a contradiction that the image of Φ is not BCA_n. The image of Φ is contained in BCA_n, so $BCA_n \setminus \mathrm{im}(\Phi)$ must be non-empty. Let τ be an automorphism in $BCA_n \setminus \mathrm{im}(\Phi)$ minimizing $c(\tau)$. Since τ is basis-conjugating, for each $1 \le r \le n$, there exists $w_r \in \mathcal{W}_n$ with $w_r x_r w_r^{-1}$ reduced and $\tau(x_r) = w_r x_r w_r^{-1}$.

For any $1 \le i, j \le n$ with $i \ne j$, it is straightforward to check that $\tau \circ \sigma_{ij}^{\pm 1}$ must be in $BCA_n \setminus \mathrm{im}(\Phi)$. By the minimality of τ, $c(\tau \circ \sigma_{ij}^{\pm 1}) \ge c(\tau)$. By lemma 1, it follows that w_j does not begin with $w_i x_i^{\pm 1}$ for all $1 \le i, j \le n$ with $i \ne j$.

Claim 1. *For any word $u \in \mathcal{W}_n$, $\ell(\tau(u)) \ge \ell(u)$.*

Proof. The result of reducing u is a word that can be written in the form $x_{r_1}^{\alpha_1} \dots x_{r_s}^{\alpha_s} \in \mathcal{W}_n$, where $r_t \ne r_{t+1}$ for each $1 \le t \le s-1$. Then,

$$\tau(u) = w_{r_1} x_{r_1}^{\alpha_1} w_{r_1}^{-1} \dots w_{r_s} x_{r_s}^{\alpha_s} w_{r_s}^{-1}.$$

For each $1 \le t \le s$, by our choice of w_{r_t}, $w_{r_t} x_{r_t} w_{r_t}^{-1}$ is reduced, which implies $w_{r_t} x_{r_t}^{\alpha_t} w_{r_t}^{-1}$ is reduced. Moreover, by the observation we made prior to the statement of claim 1, for each $1 \le t \le s-1$, w_{r_t} does not begin with $w_{r_{t+1}} x_{r_{t+1}}^{\pm 1}$ and $w_{r_{t+1}}$ does not begin with $w_{r_t} x_{r_t}^{\pm 1}$. It follows that in every possible way of reducing $\tau(u)$, for all $1 \le t \le s$, no letter in the instance of $x_{r_t}^{\alpha_t}$ in the expression for $\tau(u)$ above is deleted. Therefore, $\ell(\tau(u)) \ge \ell(u)$. □

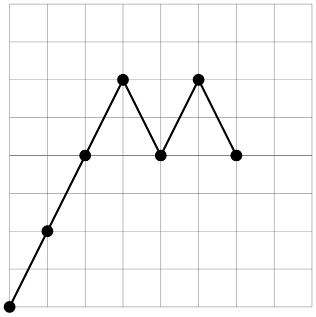
By our claim, $\ell(\tau(\tau^{-1}(x_r))) \geq \ell(\tau^{-1}(x_r))$ for each $1 \leq r \leq n$. It follows that $\tau^{-1}(x_r) = x_q$ for some $1 \leq q \leq n$. Therefore, τ acts on the standard basis $\{x_1, \ldots, x_n\}$ by permuting its elements. However, τ is basis-conjugating, so τ must be the identity map. Hence, $\tau \in \operatorname{im}(\Phi)$, which is a contradiction, thus proving that the image of Φ is indeed BCA_n.

3.4. **Injectivity of Φ.**

3.4.1. *Setup.* For a word $w = a_1 \ldots a_s \in \mathcal{W}(\Sigma_n)$, where $a_1, \ldots, a_s$ are the letters of w, and $0 \leq r \leq s$, let:

- $\phi_r(w) = a_1 \circ \ldots \circ a_r$, with $\phi_0(w)$ being the identity;
- $\phi(w) = a_1 \circ \ldots \circ a_s = \phi_s(w)$, in a slight abuse of notation;
- $c_r(w) = c(a_1 \circ a_2 \circ \ldots \circ a_r) = c(\phi_r(w))$; and
- $C(w) = (c_0(w), \ldots, c_s(w))$, which we call the *complexity sequence* of W.

We will use a visualization of the complexity sequence of w, which we call the *complexity graph* of w, consisting of the points $(r, c_r(w))$ for $r = 0, \ldots, s$ plotted, along with line segments joining consecutive points as in a line graph. An example of a complexity graph in the case where $n = 3$ and $w = \sigma_{21}^{-1}\sigma_{31}\sigma_{32}\sigma_{21}\sigma_{12}^{-1}\sigma_{32}^{-1}$, corresponding to the complexity sequence $(0, 2, 4, 6, 4, 6, 4)$, is shown below.



By definition, the complexity sequence of a word $w = a_1 \ldots a_s \in \mathcal{W}(\Sigma_n)$ always begins at 0 (i.e. $c_0(w) = 0$). A slightly more subtle property of $C(w)$ is that for any $0 \leq r \leq s-1$, $c_{r+1}(w) \neq c_r(w)$. To see this, note that $c_{r+1}(w) = c(\phi_{r+1}(w)) = c(\phi_r(w) \circ a_{r+1})$. By lemma 1, $c(\phi_r(w) \circ a_{r+1}) \neq c(\phi_r(w)) = c_r(w)$. Therefore, $c_{r+1}(w) \neq c_r(w)$, as desired. Translating to complexity graphs, this fact says that the complexity graph of w does not contain any horizontal line segments (which can be seen in the example above).

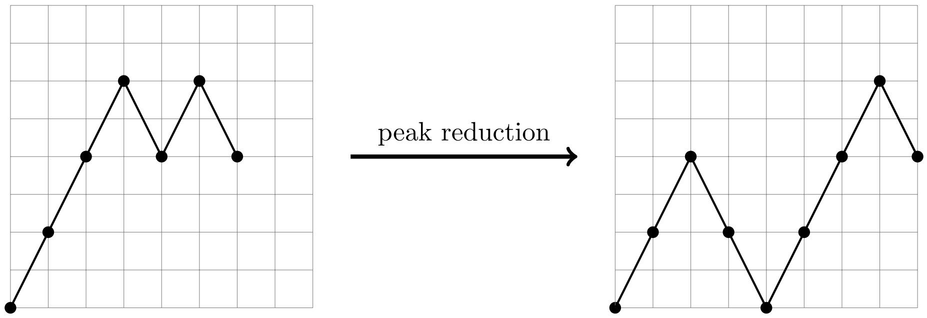
Since the complexity graph of w does not contain any horizontal line segments, each of its local maxima must be *strict* local maxima. We call a local maximum point a *peak* of the complexity graph. Translating to complexity sequences, if $c_r(w) \geq c_{r-1}(w)$ and $c_r(w) \geq c_{r+1}(w)$, then $c_r(w) > c_{r-1}(w)$ and $c_r(w) > c_{r+1}(w)$, and we call an index $1 \leq r \leq s-1$ a *peak* of $C(w)$ if $c_r(w) > c_{r-1}(w)$ and $c_r(w) > c_{r+1}(w)$.

3.4.2. *Peak Reduction.* The appearance of the word "peak" both in our preceding definitions and in the phrase "Whitehead's peak reduction lemma" is no coincidence. Suppose we keep all the definitions in the paper so far, except with

- the set of symbols Σ_n changed,
- the map $\phi : \Sigma_n \cup \Sigma_n^{-1} \to \mathrm{Aut}(\mathcal{F}_n)$ associating symbols to automorphisms changed significantly,
- the measure of complexity $c(\tau)$ changed slightly, and
- the definition of peak changed very slightly.

The details for the changes will not be specified here, but they should be clear upon consulting [HL]. In this modified context, Whitehead's peak reduction lemma translates to the statement that given a word $w = a_1 \dots a_s \in W(\Sigma_n)$ and a peak m of $C(w)$, we may replace w with another word w' such that $\phi(w') = \phi(w)$ and $C(w')$ is the same as $C(w)$, except with $c_m(w)$ replaced by a sequence of smaller numbers.

In terms of the transition from the complexity graph of w to that of w', the peak $(m, c_m(w))$ is replaced with a sequence of points lower (i.e. with smaller y-coordinates) than it. Informally, the *peak* m (or the peak $(m, c_m(w))$) is *reduced* when replacing w by w', which explains the name "peak reduction". An example of what this looks like visually is shown below, where our version of the peak reduction lemma (which we haven't stated yet) is applied to the first peak of the word from our example of a complexity graph earlier.



Returning to our context, the statement of the main lemma we use to prove the injectivity of Φ is similar to the translation of Whitehead's peak reduction lemma above. Prior to stating this lemma, we recall that the elements of PWB_n were defined to be equivalence classes of $\mathcal{F}(\Sigma_n)$, and the elements of $\mathcal{F}(\Sigma_n)$ were defined to be equivalence classes of $\mathcal{W}(\Sigma_n)$. Our convention will be to consider elements of PWB_n directly as equivalence classes of $\mathcal{W}(\Sigma_n)$.

3.4.3. *The Main Lemma.* With the convention set out in the last paragraph in mind, the statement of our version of the peak reduction lemma is:

Lemma 2. *Let $w = a_1 \dots a_s \in \mathcal{W}(\Sigma_n)$, where $a_1, \dots, a_s$ are the letters of w, and suppose m is peak of $C(w)$. Then, there exists a word $w' \in \mathcal{W}(\Sigma_n)$ such that $w' = w$*

in PWB_n and $C(w')$ is the same as $C(w)$, except with $c_m(w)$ replaced by a sequence of smaller numbers.

Proof. We begin the proof by claiming that in order to find a w' as required by the lemma, it suffices to find a word $u \in \mathcal{W}(\Sigma_n)$ meeting certain requirements.

Claim 2. *Suppose there exists a (possibly empty) word $u = b_1 \dots b_t \in \mathcal{W}(\Sigma_n)$ (t may be zero) with $a_m a_{m+1} = b_1 \dots b_t$ in PWB_n, and such that $c(\phi_{m-1}(w) \circ \phi_r(u)) < c(\phi_m(w))$ for all $1 \leq r \leq t-1$. Then, a w' satisfying the requirements of the lemma exists.*

Proof. If such a u exists, then we can replace $a_m a_{m+1}$ with u to get the word $a_1 \dots a_{m-1} u a_{m+2} \dots a_s \in \mathcal{W}(\Sigma_n)$. We let $w' = a_1 \dots a_{m-1} u a_{m+2} \dots a_s$, and show this choice of w' meets the requirements of the lemma. Clearly, $w' = w$ in PWB_n since $a_m a_{m+1} = u$ in PWB_n, so the first requirement of the lemma is met.

Since Φ is a well-defined map on PWB_n, $a_m a_{m+1}$ being equal to u in PWB_n implies that $\phi(a_m a_{m+1}) = \phi(u)$. It follows that $c_r(w) = c_{r-2+t}(w')$ for all $m+1 \leq r \leq s$. It is also clear that $c_r(w) = c_r(w')$ for all $1 \leq r \leq m-1$. For $m \leq r \leq m+t-2$, $c_r(w') = c(\phi_{m-1}(w) \circ \phi_{r-m+1}(u))$, and $c(\phi_{m-1}(w) \circ \phi_{r-m+1}(u)) < c_m(w)$ by our hypothesis on u, which implies $c_r(w') < c_m(w)$. We conclude that the sequence $C(w') = (c_0(w'), \dots, c_{s+t-2}(w'))$ is the same as $C(w)$, except with $c_m(w)$ replaced by the sequence of smaller numbers $(c_m(w'), \dots, c_{m+t-2}(w'))$. Therefore, w' meets the second requirement of the lemma as well, which proves our claim. □

Now, we prove that a u satisfying the conditions of claim 2 always exists. Our choice of u will depend on a_m and a_{m+1}. Specifically, since m is a peak for $C(w)$, $c_{m-1}(w) < c_m(w)$ and $c_{m+1}(w) < c_m(w)$. As $c_m(w) = c(\phi_m(w))$, $c_{m+1}(w) = c(\phi_{m+1}(w)) = c(\phi_m(w) \circ a_{m+1})$ and $c_{m-1}(w) = c(\phi_{m-1}(w)) = c(\phi_m(w) \circ a_m^{-1})$, we have that $c(\phi_m(w) \circ a_{m+1}) < c(\phi_m(w))$ and $c(\phi_m(w) \circ a_m^{-1}) < c(\phi_m(w))$.

Therefore, fixing a_m and a_{m+1} allows use to use the facts that $c(\phi_m(w) \circ a_{m+1}) < c(\phi_m(w))$ and $c(\phi_m(w) \circ a_m^{-1}) < c(\phi_m(w))$, deduced above, in conjunction with lemma 1 to make conclusions regarding $\phi_m(w)$. We now analyze each possible case for the identities of a_m and a_{m+1} to show that a u satisfying the conditions of claim 2 always exists. Hereafter, for $1 \leq d \leq n$, we let $w_d \in \mathcal{W}(\Sigma_n)$ be such that $w_d x_d w_d^{-1}$ is reduced and $\phi_m(w)(x_d) = w_d x_d w_d^{-1}$ (note $\phi_m(w)$ is basis-conjugating).

Case 1: $a_m = \sigma_{ij}^{\alpha}$, $a_{m+1} = \sigma_{kj}^{\beta}$, where $\alpha, \beta \in \{1, -1\}$ and $i, j, k \in \{1, \dots, n\}$ are distinct.

By lemma 1, w_j simultaneously starts with $w_i x_i^{\alpha}$ and $w_k x_k^{-\beta}$. There are two subcases to consider.

Subcase A: $\ell(w_i) < \ell(w_k)$.

Then, w_k starts $w_i x_i^{\alpha}$. Therefore, we can write $w_k = w_i x_i^{\alpha} y$ and $w_j = w_i x_i^{\alpha} y x_k^{-\beta} z$ for some words $y, z \in \mathcal{W}_n$, where $w_i x_i^{\alpha} y$ and $w_i x_i^{\alpha} y x_k^{-\beta} z$ are reduced. Let $u = b_1 b_2 b_3 b_4 = \sigma_{ik}^{-\alpha} \sigma_{kj}^{\beta} \sigma_{ij}^{\alpha} \sigma_{ik}^{\alpha}$. Note $u = a_m a_{m+1}$ in PWB_n.

We organize the information relevant to the situation in the diagram on the next page. The conventions for the diagram are stated below. The diagrams in subsequent cases should be understood as following these conventions as well.

- In the box labelled with the automorphism τ, the action of τ on the standard basis of $\mathcal{F}_n$ is described inside it, with the convention that for $1 \leq e \leq n$ and $v \in \mathcal{W}_n$, we write $x_e \longmapsto v$ if $\tau(x_e) = v x_e v^{-1}$ (for brevity).
- An arrow labelled with a symbol $g \in \Sigma_n \cup \Sigma_n^{-1}$ is drawn from the box for the automorphism τ_1 to the box for the automorphism τ_2 if $\tau_2 = \tau_1 \circ g$.

From the diagram below, it is clear that $c(\phi_{m-1}(w) \circ \phi_r(u)) < c(\phi_m(w))$ for $1 \leq r \leq 3$, so u meets the requirements of claim 2.

$$\begin{tikzcd}
& \begin{array}{|l|}\hline \phi_{m-1}(w) \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i y x_k^{-\beta} z \\ x_k \longmapsto w_i x_i^{\alpha} y \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} \arrow[dl, "\sigma_{ik}^{-\alpha}"'] \arrow[ddr, "\sigma_{ij}^{\alpha}"] & \\
\begin{array}{|l|}\hline \phi_{m-1}(w) \circ \phi_1(u) \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i y x_k^{-\beta} z \\ x_k \longmapsto w_i y \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} \arrow[d, "\sigma_{kj}^{\beta}"'] & & \\
\begin{array}{|l|}\hline \phi_{m-1}(w) \circ \phi_2(u) \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i y z \\ x_k \longmapsto w_i y \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} \arrow[d, "\sigma_{ij}^{\alpha}"'] & & \begin{array}{|l|}\hline \phi_m(w) \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i x_i^{\alpha} y x_k^{-\beta} z \\ x_k \longmapsto w_i x_i^{\alpha} y \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} \arrow[ddl, "\sigma_{kj}^{\beta}"] \\
\begin{array}{|l|}\hline \phi_{m-1}(w) \circ \phi_3(u) \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i x_i^{\alpha} y z \\ x_k \longmapsto w_i y \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} \arrow[dr, "\sigma_{ik}^{\alpha}"] & & \\
& \begin{array}{|l|}\hline \phi_{m+1}(w) \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i x_i^{\alpha} y z \\ x_k \longmapsto w_i x_i^{\alpha} y \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} &
\end{tikzcd}$$

Subcase B: $\ell(w_k) < \ell(w_i)$.

Then, w_i starts with $w_k x_k^{-\beta}$. Therefore, we can write $w_i = w_k x_k^{-\beta} y$ and $w_j = w_k x_k^{-\beta} y x_i^{\alpha} z$ for some words $y, z \in \mathcal{W}_n$, where $w_k x_k^{-\beta} y$ and $w_k x_k^{-\beta} y x_i^{\alpha} z$ are reduced. Let $u = b_1 b_2 b_3 b_4 = \sigma_{ki}^{\beta} \sigma_{kj}^{\beta} \sigma_{ij}^{\alpha} \sigma_{ki}^{-\beta}$. Note $u = a_m a_{m+1}$ in PWB_n. We organize the information relevant to the situation in the diagram below, following the same conventions as the diagram earlier. From the diagram it is clear that $c(\phi_{m-1}(w) \circ \phi_r(u)) < c(\phi_m(w))$ for $1 \leq r \leq 3$, so u meets the requirements of claim 2.

$$\begin{tikzcd}
 & \boxed{\begin{array}{l} \multicolumn{1}{c}{\phi_{m-1}(w)} \\ \hline x_i \longmapsto w_k x_k^{-\beta} y \\ x_j \longmapsto w_k x_k^{-\beta} yz \\ x_k \longmapsto w_k \\ x_d \longmapsto w_d, d \neq i, j, k \end{array}} \arrow[dl, "\sigma_{ki}^{\beta}"'] \arrow[ddr, "\sigma_{ij}^{\alpha}"] & \\
\boxed{\begin{array}{l} \multicolumn{1}{c}{\phi_{m-1}(w) \circ \phi_1(u)} \\ \hline x_i \longmapsto w_k y \\ x_j \longmapsto w_k x_k^{-\beta} yz \\ x_k \longmapsto w_k \\ x_d \longmapsto w_d, d \neq i, j, k \end{array}} \arrow[d, "\sigma_{kj}^{\beta}"'] & & \\
\boxed{\begin{array}{l} \multicolumn{1}{c}{\phi_{m-1}(w) \circ \phi_2(u)} \\ \hline x_i \longmapsto w_k y \\ x_j \longmapsto w_k yz \\ x_k \longmapsto w_k \\ x_d \longmapsto w_d, d \neq i, j, k \end{array}} \arrow[d, "\sigma_{ij}^{\alpha}"'] & & \boxed{\begin{array}{l} \multicolumn{1}{c}{\phi_m(w)} \\ \hline x_i \longmapsto w_k x_k^{-\beta} y \\ x_j \longmapsto w_k x_k^{-\beta} y x_i^{\alpha} z \\ x_k \longmapsto w_k \\ x_d \longmapsto w_d, d \neq i, j, k \end{array}} \arrow[ddl, "\sigma_{kj}^{\beta}"'] \\
\boxed{\begin{array}{l} \multicolumn{1}{c}{\phi_{m-1}(w) \circ \phi_3(u)} \\ \hline x_i \longmapsto w_k y \\ x_j \longmapsto w_k y x_i^{\alpha} z \\ x_k \longmapsto w_k \\ x_d \longmapsto w_d, d \neq i, j, k \end{array}} \arrow[dr, "\sigma_{ki}^{-\beta}"] & & \\
 & \boxed{\begin{array}{l} \multicolumn{1}{c}{\phi_{m+1}(w)} \\ \hline x_i \longmapsto w_k x_k^{-\beta} y \\ x_j \longmapsto w_k y x_i^{\alpha} z \\ x_k \longmapsto w_k \\ x_d \longmapsto w_d, d \neq i, j, k \end{array}} &
\end{tikzcd}$$

Case 2: $a_m = \sigma_{ij}^{\alpha}$, $a_{m+1} = \sigma_{jk}^{\beta}$, where $\alpha, \beta \in \{1, -1\}$ and $i, j, k \in \{1, \dots, n\}$ are distinct.

By lemma 1, w_j starts with $w_i x_i^{\alpha}$ and w_k starts with $w_j x_j^{-\beta}$. Therefore, we can write $w_j = w_i x_i^{\alpha} y$ and $w_k = w_i x_i^{\alpha} y x_j^{-\beta} z$ for some words $y, z \in \mathcal{W}_n$, where $w_i x_i^{\alpha} y$ and $w_i x_i^{\alpha} y x_j^{-\beta} z$ are reduced. Let $u = b_1 b_2 b_3 b_4 = \sigma_{ik}^{-\alpha} \sigma_{jk}^{\beta} \sigma_{ij}^{\alpha} \sigma_{ik}^{\alpha}$. Note $u = a_m a_{m+1}$ in PWB_n. From the diagram below, it is clear that $c(\phi_{m-1}(w) \circ \phi_r(u)) < c(\phi_m(w))$ for $1 \leq r \leq 3$, so u meets the requirements of claim 2.

$$\begin{tikzcd}
 & \begin{array}{|l|}\hline \phi_{m-1}(w) \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i y \\ x_k \longmapsto w_i x_i^{\alpha} y x_j^{-\beta} z \\ x_d \longmapsto w_d, d \neq i,j,k \\ \hline \end{array} \arrow[dl, "\sigma_{ik}^{-\alpha}"'] \arrow[ddr, "\sigma_{ij}^{\alpha}"] & \\
\begin{array}{|l|}\hline \phi_{m-1}(w) \circ \phi_1(u) \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i y \\ x_k \longmapsto w_i y x_j^{-\beta} z \\ x_d \longmapsto w_d, d \neq i,j,k \\ \hline \end{array} \arrow[d, "\sigma_{jk}^{\beta}"'] & & \\
\begin{array}{|l|}\hline \phi_{m-1}(w) \circ \phi_2(u) \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i y \\ x_k \longmapsto w_i y z \\ x_d \longmapsto w_d, d \neq i,j,k \\ \hline \end{array} \arrow[d, "\sigma_{ij}^{\alpha}"'] & & \begin{array}{|l|}\hline \phi_m(w) \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i x_i^{\alpha} y \\ x_k \longmapsto w_i x_i^{\alpha} y x_j^{-\beta} z \\ x_d \longmapsto w_d, d \neq i,j,k \\ \hline \end{array} \arrow[ddl, "\sigma_{jk}^{\beta}"] \\
\begin{array}{|l|}\hline \phi_{m-1}(w) \circ \phi_3(u) \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i x_i^{\alpha} y \\ x_k \longmapsto w_i y z \\ x_d \longmapsto w_d, d \neq i,j,k \\ \hline \end{array} \arrow[dr, "\sigma_{ik}^{\alpha}"] & & \\
 & \begin{array}{|l|}\hline \phi_{m+1}(w) \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i x_i^{\alpha} y \\ x_k \longmapsto w_i x_i^{\alpha} y z \\ x_d \longmapsto w_d, d \neq i,j,k \\ \hline \end{array} &
\end{tikzcd}$$

Case 3: $a_m = \sigma_{ji}^{\alpha}$, $a_{m+1} = \sigma_{kj}^{\beta}$, where $\alpha, \beta \in \{1, -1\}$ and $i, j, k \in \{1, \dots, n\}$ are distinct.

By lemma 1, w_i starts with $w_j x_j^{\alpha}$ and w_j starts with $w_k x_k^{-\beta}$. Therefore, we can write $w_j = w_k x_k^{-\beta} y$ and $w_i = w_k x_k^{-\beta} y x_j^{\alpha} z$ for some words $y, z \in \mathcal{W}_n$, where $w_k x_k^{-\beta} y$ and $w_k x_k^{-\beta} y x_j^{\alpha} z$ are reduced. Let $u = b_1 b_2 b_3 b_4 = \sigma_{ki}^{\beta} \sigma_{kj}^{\beta} \sigma_{ji}^{\alpha} \sigma_{ki}^{-\beta}$. Note $u = a_m a_{m+1}$ in PWB_n. From the diagram below, it is clear that $c(\phi_{m-1}(w) \circ \phi_r(u)) < c(\phi_m(w))$ for $1 \leq r \leq 3$, so u meets the requirements of claim 2.

$$\begin{tikzcd}
 & \begin{array}{|l|}\hline \phi_{m-1}(w) \\ \hline x_i \longmapsto w_k x_k^{-\beta} yz \\ x_j \longmapsto w_k x_k^{-\beta} y \\ x_k \longmapsto w_k \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} \arrow[dl, "\sigma_{ki}^{\beta}"'] \arrow[ddr, "\sigma_{ji}^{\alpha}"] & \\
\begin{array}{|l|}\hline \phi_{m-1}(w) \circ \phi_1(u) \\ \hline x_i \longmapsto w_k yz \\ x_j \longmapsto w_k x_k^{-\beta} y \\ x_k \longmapsto w_k \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} \arrow[d, "\sigma_{kj}^{\beta}"'] & & \\
\begin{array}{|l|}\hline \phi_{m-1}(w) \circ \phi_2(u) \\ \hline x_i \longmapsto w_k yz \\ x_j \longmapsto w_k y \\ x_k \longmapsto w_k \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} \arrow[d, "\sigma_{ji}^{\alpha}"'] & & \begin{array}{|l|}\hline \phi_m(w) \\ \hline x_i \longmapsto w_k x_k^{-\beta} y x_j^{\alpha} z \\ x_j \longmapsto w_k x_k^{-\beta} y \\ x_k \longmapsto w_k \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} \arrow[ddl, "\sigma_{kj}^{\beta}"] \\
\begin{array}{|l|}\hline \phi_{m-1}(w) \circ \phi_3(u) \\ \hline x_i \longmapsto w_k y x_j^{\alpha} z \\ x_j \longmapsto w_k y \\ x_k \longmapsto w_k \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} \arrow[dr, "\sigma_{ki}^{-\beta}"] & & \\
 & \begin{array}{|l|}\hline \phi_{m+1}(w) \\ \hline x_i \longmapsto w_k x_k^{-\beta} y x_j^{\alpha} z \\ x_j \longmapsto w_k y \\ x_k \longmapsto w_k \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} &
\end{tikzcd}$$

Case 4: $a_m = \sigma_{ji}^{\alpha}$, $a_{m+1} = \sigma_{jk}^{\beta}$, where $\alpha, \beta \in \{1, -1\}$ and $i, j, k \in \{1, \ldots, n\}$ are distinct.

By lemma 1, w_i starts with $w_j x_j^{\alpha}$ and w_k starts with $w_j x_j^{-\beta}$. Therefore, can write $w_i = w_j x_j^{\alpha} y$ and $w_k = w_j x_j^{-\beta} z$ for some words $y, z \in \mathcal{W}_n$, where $w_j x_j^{\alpha} y$ and $w_j x_j^{-\beta} z$ are reduced. Let $u = b_1 b_2 = \sigma_{jk}^{\beta} \sigma_{ji}^{\alpha}$. Note $u = a_m a_{m+1}$ in PWB_n. From the diagram below, it is clear that $c(\phi_{m-1}(w) \circ \phi_1(u)) < c(\phi_m(w))$, so u meets the requirements of claim 2.

$$\begin{tikzcd}
 & \begin{array}{|l|}\hline \multicolumn{1}{|c|}{\phi_{m-1}(w)} \\ \hline x_i \longmapsto w_j y \\ x_j \longmapsto w_j \\ x_k \longmapsto w_j x_j^{-\beta} z \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} \arrow[dl, "\sigma_{jk}^{\beta}"'] \arrow[dr, "\sigma_{ji}^{\alpha}"] & \\
\begin{array}{|l|}\hline \multicolumn{1}{|c|}{\phi_{m-1}(w) \circ \phi_1(u)} \\ \hline x_i \longmapsto w_j y \\ x_j \longmapsto w_j \\ x_k \longmapsto w_j z \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} \arrow[dr, "\sigma_{ji}^{\alpha}"'] & & \begin{array}{|l|}\hline \multicolumn{1}{|c|}{\phi_m(w)} \\ \hline x_i \longmapsto w_j x_j^{\alpha} y \\ x_j \longmapsto w_j \\ x_k \longmapsto w_j x_j^{-\beta} z \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} \arrow[dl, "\sigma_{jk}^{\beta}"] \\
 & \begin{array}{|l|}\hline \multicolumn{1}{|c|}{\phi_{m+1}(w)} \\ \hline x_i \longmapsto w_j x_j^{\alpha} y \\ x_j \longmapsto w_j \\ x_k \longmapsto w_j z \\ x_d \longmapsto w_d, d \neq i, j, k \\ \hline \end{array} &
\end{tikzcd}$$

Case 5: $a_m = \sigma_{ij}^{\alpha}$, $a_{m+1} = \sigma_{pq}^{\beta}$, where $\alpha, \beta \in \{1, -1\}$ and $i, j, p, q \in \{1, \ldots, n\}$ are distinct.

By lemma 1, w_j starts with $w_i x_i^{\alpha}$ and w_q starts with $w_p x_p^{-\beta}$. Therefore, we can write $w_j = w_i x_i^{\alpha} y$ and $w_q = w_p x_p^{-\beta} z$ for some words $y, z \in \mathcal{W}_n$, where $w_i x_i^{\alpha} y$ and $w_p x_p^{-\beta} z$ are reduced. Let $u = b_1 b_2 = \sigma_{pq}^{\beta} \sigma_{ij}^{\alpha}$. Note $u = a_m a_{m+1}$ in PWB_n. From the diagram below, it is clear that $c(\phi_{m-1}(w) \circ \phi_1(u)) < c(\phi_m(w))$, so u meets the requirements of claim 2.

$$\begin{tikzcd}
 & \boxed{\begin{array}{l} \multicolumn{1}{c}{\phi_{m-1}(w)} \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i y \\ x_p \longmapsto w_p \\ w_q \longmapsto w_p x_p^{-\beta} z \\ x_d \longmapsto w_d, d \neq i, j, p, q \end{array}} \arrow[dl, "\sigma_{pq}^{\beta}"'] \arrow[dr, "\sigma_{ij}^{\alpha}"] & \\
\boxed{\begin{array}{l} \multicolumn{1}{c}{\phi_{m-1}(w) \circ \phi_1(u)} \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i y \\ x_p \longmapsto w_p \\ w_q \longmapsto w_p z \\ x_d \longmapsto w_d, d \neq i, j, p, q \end{array}} \arrow[dr, "\sigma_{ij}^{\alpha}"'] & & \boxed{\begin{array}{l} \multicolumn{1}{c}{\phi_m(w)} \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i x_i^{\alpha} y \\ x_p \longmapsto w_p \\ w_q \longmapsto w_p x_p^{-\beta} z \\ x_d \longmapsto w_d, d \neq i, j, p, q \end{array}} \arrow[dl, "\sigma_{pq}^{\beta}"] \\
 & \boxed{\begin{array}{l} \multicolumn{1}{c}{\phi_{m+1}(w)} \\ \hline x_i \longmapsto w_i \\ x_j \longmapsto w_i x_i^{\alpha} y \\ x_p \longmapsto w_p \\ w_q \longmapsto w_p z \\ x_d \longmapsto w_d, d \neq i, j, p, q \end{array}} &
\end{tikzcd}$$

Case 6: $a_m = \sigma_{ij}^{\alpha}$, $a_{m+1} = \sigma_{ij}^{\beta}$, where $\alpha, \beta \in \{1, -1\}$ and $i, j \in \{1, \ldots, n\}$ are distinct.

By lemma 1, w_j simultaneously begins with $w_i x_i^{\alpha}$ and $w_i x_i^{-\beta}$. Therefore, $\alpha = -\beta$, which implies $a_m a_{m+1} = 1$ in PWB_n, so we can take u to be the empty word, and it is easy to see that this choice of u satisfies the conditions of claim 2.

Case 7: $a_m = \sigma_{ij}^{\alpha}$, $a_{m+1} = \sigma_{ji}^{\beta}$, where $\alpha, \beta \in \{1, -1\}$ and $i, j \in \{1, \ldots, n\}$ are distinct.

By lemma 1, w_j starts with $w_i x_i^{\alpha}$ and w_i starts with $w_j x_i^{-\beta}$. Since we cannot simultaneously have w_j starting with w_i and w_i starting with w_j, this case is impossible.

All possibilities for a_m and a_{m+1} are exhausted in the analysis above, and in each valid case we showed that a u satisfying the conditions of claim 2 exists. By claim 2, the proof of the lemma is complete. □

3.4.4. *Conclusion.* We now apply lemma 2 in order to show that Φ is injective. We use the phrase *applying peak reduction* to refer to the process of using the algorithm implicit in the proof of lemma 2 to generate a word $w' \in \mathcal{W}(\Sigma_n)$, given a word $w = a_1 \ldots a_s \in \mathcal{W}(\Sigma_n)$ and a peak r of $C(w)$.

Now, suppose for a contradiction that Φ is not injective. It follows that there exists a word $w \in \mathcal{W}(\Sigma_n)$ with $\phi(w) = 1$ and $w \neq 1$ in PWB_n. Then, the complexity sequence of w begins and ends with 0. Now, let v be the result of repeatedly applying peak reduction to w until it is no longer possible to. Since applying peak reduction to a word preserves its equivalence class in PWB_n, we have that $v = w$ in PWB_n.

Since Φ is a well-defined map on PWB_n, it follows that $\phi(v) = \phi(w)$, and so the complexity sequence of v also begins and ends with 0. Since it is not possible to apply peak reduction to v, the complexity graph/sequence of v must not have any peaks. Therefore, the complexity graph of v must start at $(0, 0)$, end at some $(s', 0)$, and not have any horizontal line segments. This is only possible if v is the empty word. Hence, $w = v = 1$ in PWB_n, which is a contradiction, thus showing that Φ is indeed injective.

We have shown that the image of Φ is BCA_n and that Φ is injective. Therefore, the proof of McCool's theorem is complete. □

4. Acknowledgements

I would like to thank my supervisor, Dror Bar-Natan, for suggesting the problem of simplifying the proof of McCool's theorem, for the numerous discussions that led to the creation of this paper, and for his recommendations on writing this paper.

References

[A] Emil Artin, *Theory of Braids*, Annals of Mathematics (2) **48** (1947), 101–126.

[BNDV] Dror Bar-Natan, Zsuzsanna Dancso, and Roland van der Veen, *Over then Under Tangles* (2020), arXiv:2007.09828.

[FRR] Roger Fenn, Richard Rimanyi, and Colin Rourke, *The braid-permutation group*, Topology **36** (1997), 123–135.

[HL] P. J. Higgins and R. C. Lyndon, *Equivalence of Elements Under Automorphisms of a Free Group*, Journal of the London Mathematical Society (2) **8** (1974), 254–258.

[M1] James McCool, *Some finitely presented subgroups of the automorphism group of a free group*, Journal of Algebra **35** (1975), 205–213.

[M2] ———, *On Basis-Conjguating Automorphisms of Free Groups*, Canadian Journal of Mathematics **38** (1986), 1525–1529.

[R] Elvira Rapaport, *On free groups and their automorphisms*, Acta Mathematica **99** (1958), 138–163.

[W] J. H. C. Whitehead, *On Equivalent Sets of Elements in a Free Group*, Annals of Mathematics (2) **37** (1936), 782–800.

Department of Mathematics, University of Toronto, Toronto Ontario M5S 2E4, Canada

Email address: `sina.abbasi@mail.utoronto.ca`