COMBINATORIAL AND METRIC PROPERTIES
OF
THOMPSON’S GROUP $T$

JOSÉ BURILLO, SEAN CLEARY, MELANIE STEIN, AND JENNIFER TABACK

Abstract. We discuss metric and combinatorial properties of Thompson’s group $T$, such as the normal forms for elements and uniqueness of tree pair diagrams. We relate these properties to those of Thompson’s group $F$ when possible, and highlight combinatorial differences between the two groups. We define a set of unique normal forms for elements of $T$ arising from minimal factorizations of elements into convenient pieces. We show that the number of carets in a reduced representative of $T$ estimates the word length, that $F$ is undistorted in $T$, and that cyclic subgroups of $T$ are undistorted. We show that every element of $T$ has a power which is conjugate to an element of $F$ and describe how to recognize torsion elements in $T$.

1. Introduction

Thompson’s groups $F$, $T$ and $V$ are a remarkable family of infinite, finitely-presentable groups studied for their own interesting properties as well as for their connections with questions in logic, homotopy theory and measure theory of discrete groups.

Cannon, Floyd and Parry give an excellent introduction to these groups in [5]. These three groups can be studied either algebraically, analytically or geometrically. Algebraically, each has both finite and infinite presentations. Geometrically, an element in each group can be viewed as a tree pair diagram; that is, as a pair of finite binary rooted trees with the same number of leaves, with a numbering system pairing the leaves in the two trees. Analytically, an element of each group can be viewed as a self map of the unit interval:

- in $F$ as a piecewise linear homeomorphism,
- in $T$ as a homeomorphism of the unit interval with the endpoints identified, and thus of $S^1$.
in $V$ as a right-continuous bijection which is locally orientation preserving.

Thompson’s group $F$ in particular has been studied extensively. The group $F$ has a standard infinite presentation in which every element has a unique normal form, and a standard two-generator finite presentation. Fordham [8] presented a method of computing the word length of $w \in F$ with respect to the standard finite generating set directly from the tree pair diagram representing $w$. Regarding $F$ as a diagram group, Guba [10] also obtained an effective geometric method for computing the word metric with respect to the standard finite generating set. Belk and Brown [1] have similar results which arise from viewing elements of $F$ as forest diagrams.

In this paper, we discuss analogues for $T$ of some properties of $F$, such as normal forms for elements. We consider metrically how $F$ is contained as a subgroup of $T$, and show that the number of carets in a reduced tree pair diagram representing $w \in T$ estimates the word length of $w$ with respect to a particular generating set. We show that cyclic subgroups of $T$ are undistorted and that every element of $T$ has a power which is conjugate to an element of $F$. The groups $T$ and $V$, unlike $F$, contain torsion elements, and we describe how to recognize these torsion elements from their tree pair diagrams.

2. Background on Thompson’s Groups $F$ and $T$

2.1. Presentations and tree pair diagrams. Thompson’s groups $F$ and $T$ both have representations as groups of piecewise-linear homeomorphisms. The group $F$ is the group of orientation-preserving homeomorphisms of the interval $[0,1]$, where each homeomorphism is required to have only finitely many discontinuities of slope, called breakpoints, have slopes be powers of two and have the coordinates of the breakpoints all lie in the set of dyadic rationals. Similarly, the group $T$ consists of orientation-preserving homeomorphisms of the circle $S^1$ satisfying the same conditions where we represent the circle $S^1$ as the unit interval $[0,1]$ with the two endpoints identified.

Cannon, Floyd and Parry give an excellent introduction to Thompson’s groups $F$, $T$ and $V$ in [5]. We refer the reader to this paper for full details on results mentioned in this section. Both $F$ and $T$ can be studied either through finite or infinite presentations. With respect to the infinite presentation

$$\langle x_i, i \geq 0 \mid x_j x_i = x_i x_{j+1}, i < j \rangle$$

for $F$, group elements have simple normal forms which are unique. It is easy to see that $F$ can be generated by $x_0$ and $x_1$, which form the standard finite generating set for $F$, and yield the finite presentation
The group $T$ also has both a finite and an infinite presentation. The infinite presentation is given by two families of generators, $\{x_i, i \geq 0\}$, the same generators as in the infinite presentation of $F$, a family $\{c_i, i \geq 0\}$ of torsion elements, and the following relators:

1. $x_j x_i = x_i x_{j+1}$, if $i < j$
2. $c_n x_k = x_{k+1} c_{n+1}$, if $k < n$
3. $c_n = x_0 c_{n+1}^2$
4. $c_n x_n = c_{n+1}$.

Using the first three relators, we see that only the generators $x_0, x_1$ and $c_1$ are required to generate the group, since the other generators can be obtained from these three. In the following, we will use $c$ to denote the generator $c_1$. The group $T$ is finitely presented using the following relators, both with respect to the infinite generating set and the finite generating set $\{x_0, x_1, c\}$:

1. $[x_0 x_1^{-1}, x_0^{-1} x_1 x_0]$  
2. $[x_0 x_1^{-1}, x_0^{-2} x_1 x_0^2]$  
3. $c_2 x_1 = x_2 c_3$, (that is, $x_1^{-1} c x_0 = c x_1$)  
4. $c_1 = x_0 c_2^2$, (that is, $c = x_0 (x_1^{-1} c x_0)^2$)  
5. $c_1 x_1 = c_2$, (that is, $x_1^{-1} c x_0 x_1 = x_0^{-1} x_1 x_0 x_1^{-2} c x_0^2$)  
6. $c^3 = 1$.

As with Thompson’s group $F$, we will frequently work with the more convenient infinite set of generators when constructing normal forms for elements and performing computations in the group. We will need to express elements with respect to the finite generating set when discussing word length.

A convenient representation for an element $w$ in $F$ or $T$ is a tree pair diagram, as discussed in [5]. A tree pair diagram is a pair of finite rooted binary trees with the same number of vertices, together with a numbering of the valence one vertices. A node of the tree together with its two downward directed edges is called a caret. Valence one vertices of these trees are called exposed leaves. In $F$, we insist that both leaf numberings begin at 0 and increase from left to right. In $T$, the numberings need only increase cyclically from left to right.

The left side of the tree consists of the root caret, and all carets connected to the root by a path of left edges; the right side of the tree is defined analogously. A caret is called a left caret if its left leaf lies on the left side of the tree. A caret is called a right caret if it is not the root caret and its right leaf lies on the right side of the tree. All other carets are called interior.
A caret is called exposed if it contains two exposed leaves. We write \( w = (T_-, T_+) \) to express \( w \) as a tree pair diagram, and refer to \( T_- \) as the source tree and \( T_+ \) as the target tree. Such a tree pair diagram is not unique. There are many possible diagrams representing a given element. We can choose the cyclic ordering for elements of \( T \) to always begin with 0 on the leftmost leaf of the source tree, and furthermore we impose a natural reduction condition: if \( w = (T_-, T_+) \) and both trees contain a caret with two exposed leaves numbered \( n \) and \( n + 1 \), then we remove these carets and renumber the leaves, thus forming a representative for \( w \) with fewer carets and leaves. A tree pair diagram which admits no such reductions is called a reduced tree pair diagram, and any element of \( F \) is represented by a unique reduced tree pair diagram. When we write \( w = (T_-, T_+) \) below, we are assuming that the tree pair diagram is reduced unless otherwise specified.

When \( w \in F \) or \( w \in T \), we denote the number of carets in either tree of a tree pair diagram representing \( w \) by \( N(w) \). When \( p \) is a word in the generators of \( F \) or \( T \), then \( p \) represents an element \( w \) in either \( F \) or \( T \), and we write \( N(p) \) interchangeably with \( N(w) \).

If \( w = (T_-, T_+) \in F \), then the leaves in both trees are numbered from left to right, beginning with zero. In this case, the subdivisions of the interval are paired in increasing order, so that the intervals with zero as their left endpoint are paired, and the intervals with one as their right endpoint are paired. We may omit leaf numberings for elements of \( F \) for brevity without any ambiguity. If \( w \in T \), then \( w \) corresponds to a homeomorphism of \( S^1 \) rather than \([0, 1] \). In elements of \( T \), we can omit most leaf numbers for brevity by adopting the following convention from [5]: we understand the leaves in the source tree to be numbered from 0 to \( n \) beginning with the leftmost leaf, and indicate by a circle or zero which of the leaves in the target tree is paired with the first leaf in the source tree. Other leaf numberings can be deduced from this single mark using the cyclic order. For this reason, we often refer to tree pair diagrams representing elements of \( T \) as marked tree pair diagrams.

For example, the element \( c \) corresponds to the homeomorphism of \( S^1 \) given by

\[
c(t) = \begin{cases} 
\frac{1}{2} t + \frac{3}{4} & \text{if } 0 \leq t < \frac{1}{4} \\
2t - 1 & \text{if } \frac{1}{4} \leq t < \frac{3}{4} \\
t - \frac{1}{4} & \text{if } \frac{3}{4} \leq t \leq 1 
\end{cases}
\]

and has the marked tree pair diagram given in Figure 1.

2.2. Group Multiplication in \( F \) and \( T \). Group multiplication in \( F \) and \( T \) corresponds to composition of homeomorphisms, which we can interpret on the level of tree pair diagrams as well. First, we consider \( u, v \in F \), where
Figure 1. The tree pair diagram for the generator $c$ in $T$.

Figure 2. The tree pair diagram for sample elements $u$ and $v$ in $T$.

$u = (T_-, T_+)$ and $v = (S_-, S_+)$. To compute the tree pair diagram corresponding to the product $vu$, we create unreduced representatives $(T'_-, T'_+)$ and $(S'_-, S'_+)$ of the two elements in which $T'_+ = S'_-$. Then the product is represented by the possibly unreduced tree pair diagram $(T'_-, S'_+)$. To multiply tree pair diagrams representing elements of $T$ we follow a similar procedure. We let $u, v \in T$, where $u = (T_-, T_+)$ and $v = (S_-, S_+)$. To compute the tree pair diagram corresponding to the product $vu$, we create unreduced representatives $(T'_-, T'_+)$ and $(S'_-, S'_+)$ of the two elements in which $T'_+ = S'_-$ as trees. The product $vu$ will be represented by the pair $(T'_-, S'_+)$ of trees. To decide which leaf in $S'_+$ to mark with the zero, we just note that it should be the leaf which is mapped onto by the zero leaf in $T'_-$. To identify this leaf, we find the zero leaf in $T'_+$. Since $T'_+ = S'_-$ as trees, this leaf viewed as a leaf in $S'_-$ will be labelled $m$. Then the leaf labelled $m$ in $S'_+$ will be the new zero leaf in the tree pair diagram $(T'_-, S'_+)$ for $vu$. Alternately, we can follow the composition in both pairs of trees to see how the leaves map to each other. This constructed tree pair diagram will represent $vu$ and is not necessarily reduced. For an example of this multiplication, see Figures 2, 3 and 4.

3. Words and diagrams

3.1. Normal forms and tree pair diagrams in $F$. With respect to the infinite presentation for $F$ given above, every element of $F$ has a unique normal form. To describe these, we first observe that any $w$ can be written in the form

$$w = x_{i1}^{r_1} x_{i2}^{r_2} \cdots x_{i_k}^{r_k} x_{j_1}^{-s_1} x_{j_2}^{-s_2} \cdots x_{j_l}^{-s_l}$$
where $r_i, s_i > 0, 0 \leq i_1 < i_2 \ldots < i_k$ and $0 \leq j_1 < j_2 \ldots < j_l$. However, this expression is not unique. Uniqueness is guaranteed by the addition of the following condition: when both $x_i$ and $x_i^{-1}$ occur in the expression, so does $x_{i+1}$ or $x_{i-1}$, as discussed by Brown and Geoghegan [2]. When we refer to elements of $F$ in normal form, we mean this unique normal form.

If the normal form for $w \in F$ contains no generators with negative exponents, we refer to $w$ as a *positive word* and similarly, we say a normal form is a *negative word* if there are no generators with positive exponents.

We call any word which has the form

$$w = x_{i_1}^{r_1} x_{i_2}^{r_2} \ldots x_{i_k}^{r_k} x_{j_1}^{-s_1} \ldots x_{j_l}^{-s_l}$$

where $r_i, s_i > 0, 0 \leq i_1 < i_2 \ldots < i_k$ and $0 \leq j_1 < j_2 \ldots < j_l$, a word in *pq form*, where $p$ is the positive part of the normal form and $q$ the negative part. The normal form for an element of $F$ is the shortest word among all words in *pq form* representing the given element.
To any (not necessarily reduced) tree pair diagram \((T_-, T_+\)) for an element of \(F\) we may associate a word in \(pq\) form representing the element, using the leaf exponents in the target and source trees. When the leaves of a finite rooted binary tree are numbered from left to right, beginning with zero, the leaf exponent of leaf \(k\) is the integer length of the longest string of left edges of carets which originates at leaf \(k\) and does not reach the right side of the tree. A tree pair diagram then gives the word

\[x_{i_1}^r x_{i_2}^r \ldots x_{i_m}^r x_{j_1}^{-s_1} \ldots x_{j_2}^{-s_2} x_{j_3}^{-s_3} \ldots \]

precisely when leaf \(i_k\) in \(T_+\) has exponent \(r_k\), leaf \(j_k\) in \(T_-\) has leaf exponent \(s_k\), and generators which do not appear in the word correspond to leaves with exponent zero. We think of this word as the \(pq\) factorization of the element given by the particular tree pair diagram. On the other hand, any word in \(pq\) form can be translated into a tree pair diagram. Furthermore, under this correspondence for \(F\), reduced tree pair diagrams correspond exactly to normal forms. For examples of this correspondence, see [5, 6, 7].

We observe that if an exposed caret has leaves numbered \(i\) and \(i+1\), then leaf \(i+1\) must have leaf exponent zero, since it is a right leaf. If both trees in a tree pair diagram have exposed carets with leaves numbered \(i\) and \(i+1\), then the corresponding normal form, computed via leaf exponents, contains the generators \(x_i\) to both positive and negative indices, but no instances of the generator \(x_{i+1}\). This is precisely the situation when the normal form can be reduced by a relator of \(F\). Thus the condition that the normal form is unique is exactly the condition that the tree pair diagram is reduced. This correspondence will be extended to elements of \(T\) in the next section.

3.2. Tree pair diagrams for elements of \(T\). We now discuss the relationship between words in \(T\) and tree pair diagrams. The relationship is more complicated in \(T\) than it is in \(F\). The representation of elements of \(T\) by marked tree pair diagrams suggests a way to decompose an element of \(T\) into a product of three elements: the positive and negative parts together with a torsion part in the middle, as described in [5].

**Definition 3.1.** Let the marked tree pair diagram \((T_-, T_+)\) represent \(g \in T\). If \(T_-\) and \(T_+\) each have \(i+1\) carets, then we let \(R\) be the all-right tree which has \(i+1\) carets, all of which lie on the right side of the tree. We can write \(g\) as a product of:

1. \(a \in F\) with tree pair diagram \((T_-, R)\) and \(a\) has negative normal form \(u\),
2. a cyclic permutation \(c_j^i\) for some \(i\) where \(1 \leq j \leq i+1\) (which permutes the leaf numbering in \(R\)), and
Figure 5. Three tree pair diagrams representing the word $x_1x_2c_5x_2^{-2}x_1^{-1}x_0^{-2}$ factorized as $pcq$.

(3) $b \in F$ represented by $(R, T_+)$, where $b$ has positive normal form $v$, ignoring the leaf numbering on $T_+$.

Then the word $w = vc_i^ju$ is called the $pcq$ factorization of $g$ associated to the marked tree pair diagram $(T_-, T_+)$. In the special case where $g \in F \subset T$, the $pcq$ factorization will just be the usual $pq$ factorization, as we consider the $c$ part of the word to be empty.

Figure 5 illustrates an example of an element of $T$ decomposed in this way.

The following theorem follows from the existence of these decompositions, and an algebraic proof of this result is found in [5].

**Theorem 3.2** ([5], Theorem 5.7). Any element $x \in T$ admits an expression of the form

$$x^{r_1}x^{r_2} \cdots x^{r_n} c^{j_1}x^{j_2-n} \cdots x^{j_m-n}$$

where $0 \leq i_1 < i_2 < \cdots < i_n$ and $0 \leq j_1 < j_2 < \cdots < j_m$ and either $1 \leq j < i + 2$ or $c_i^j$ is not present.

We refer to any word satisfying the hypotheses of Theorem 3.2 as a word in $pcq$ form for an element of $T$ (just as words of this form with no $c_i^j$ term are called words in $pq$ form in the group $F$). Neither proof of the existence of $pcq$ forms gives an easy explicit method for transforming a general word in the generators $x_i^{\pm1}, c_i$ into $pcq$ form without resorting to drawing tree pair diagrams, so we will outline such a method below. We recall that the four types of relators we are using in $T$ are:

1. $x_jx_i = x_ix_{j+1}$, if $i < j$
2. $c_nx_k = x_{k+1}c_{n+1}$, if $k < n$
3. $c_n = x_0c_n^{m+1}$
4. $c_n = c_{n+1}$

**Lemma 3.3** (Pumping Lemma). The generators $x_i$ and $c_j$ of $T$ satisfy

$$c_n^m = x_{m-1}c_{n+1}^{m+1} \quad \text{if} \quad 1 \leq m < n + 2.$$ 

**Proof.** This follows immediately from the relators of the types (2) and (3). $\square$
We consider a word $w \in T$ written in the generators $\{x_i, c_j\}$. To put $w$ into $pcq$ form, we must move all the positive powers of the $x_i$ to the left and all negative powers to the right, leaving only generators $c_i$ in the middle. We can use relators of types (1), (2) and (4) to accomplish that in the following way. If there is a $c_n$ followed by an $x_k$, we apply a relator of type (2) to switch them if $n > k$. If $n \leq k$, then we use the Pumping Lemma to increase the index of the $c_n$ until it is large enough, and then we use the equation

$$c_k^m x_k = c_k^{m-1} c_{k+1} = x_{m-2} c_{k+1}^{m+1},$$

where the first equality is a relator of type (4) and the second equality follows from the Pumping Lemma. The same procedure can be used for negative powers of the $x_i$ by taking inverses in the relators and the Pumping Lemma. During this process, if two generators of the type $x_n$ need to be moved past each other, we use the relators of type (1).

The result of this process is a product of a positive word in $F$, several powers of different elements of the form $c_j$, and a negative word in $F$. To combine a product of several $c_j$ into a single element $c_k$, we use the Pumping Lemma. We can always combine $c_i c_j$ to obtain an expression with a single $c_k$ by raising the lower index via the Pumping Lemma, at the cost of potentially accumulating some positive powers of $x_i$ generators at the front of the expression or negative powers of $x_i$ at the back of the expression. For example,

$$c_2^3 c_5 = x_1 c_2^3 c_5 = x_1 x_2 c_2^3 c_5 = x_1 x_2 c_2^3.$$ 

Once the resulting expression contains a single $c_i$ generator, we can again use the relators of type (1), the relators in $F$, to arrange the positive and negative words so that the generators have the appropriate increasing or decreasing order by index.

The relationship between words in $pq$ form and tree pair diagrams in $F$ is different than the relationship between $pcq$ forms and tree pair diagrams in $T$. In $F$, every tree pair diagram has a $pq$ factorization associated to it, and any word in $pq$ form is in fact the $pq$ factorization associated to a (not necessarily unique) tree pair diagram. Given any word in $pq$ form, then we can form a tree pair diagram for this element as follows. We consider reduced tree pair diagrams for $p$ and $q$, and construct a tree pair diagram for the product $pq$ as described in Section 2.2. The middle trees of the four trees involved in the product are all-right trees. The all-right trees in this decomposition may not have the same number of carets, so in forming the diagram for $pq$ we simply enlarge the smaller of the two of these all-right trees (as well as the other tree in that diagram). Since only right carets are ever added during this process, all of whose leaves have leaf exponent zero,
this results in a tree pair diagram whose \( pq \) factorization is precisely the word \( pq \) we began with.

In \( T \), the correspondence between \( pcq \) factorizations and general \( pcq \) words is not so straightforward. Although each tree pair diagram has a \( pcq \) factorization associated to it, general words in algebraic \( pcq \) form are not always the \( pcq \) factorizations associated to a tree pair diagram. The difficulty arises when the tree pair diagram for \( c \) does not have as many carets as those for \( p \) or \( q \), as adding right carets to enlarge \( c \) appropriately necessitates adding generators to the normal forms for \( p \) and \( q \), so the tree pair diagram one obtains by multiplying as in \( F \) will not necessarily have the original word as its factorization. For example, the word \( x_1c_1 \) is in algebraic \( pcq \) form, yet it is not the \( pcq \) factorization associated to some tree pair diagram. There is a different representative for this element of \( T \) which is the \( pcq \) factorization associated to the reduced tree pair diagram for this group element: \( x_1c_2x_1^{-1} \). We prefer to work with words which are \( pcq \) factorizations associated to tree pair diagrams, which will lead us to unique normal forms.

We can algebraically characterize the words of type \( pcq \) which are \( pcq \) factorizations associated to tree pair diagrams. The important condition is that the reduced tree pair diagram for \( c \) should have at least as many carets as those for \( p \) and \( q \). We say that words in \( T \) with this property satisfy the factorization condition.

**Theorem 3.4.** For elements in \( T \setminus F \), the word

\[
x_{i_1}^{r_1}x_{i_2}^{r_2}\cdots x_{i_n}^{r_n}c_{j_1}^{s_1}x_{j_2}^{-s_2}\cdots x_{j_m}^{-s_m},
\]

where \( i_1 < i_2 < \cdots < i_n, j_1 < j_2 < \cdots < j_m, \) and \( 1 \leq j < i + 2 \), is the \( pcq \) factorization associated to a tree pair diagram if and only if the number of carets in the reduced tree pair diagram for \( c \) is greater than or equal to the number of carets in the reduced tree pair diagram for either of the words \( x_{i_1}^{r_1}x_{i_2}^{r_2}\cdots x_{i_n}^{r_n} \) or \( x_{j_1}^{-s_1}x_{j_2}^{-s_2}\cdots x_{j_m}^{-s_m} \) in \( F \).

**Proof.** Given a tree pair diagram, by construction, the \( pcq \) factorization associated to it satisfies the factorization condition. Given a word that satisfies the factorization condition, we can easily construct the corresponding tree pair diagram as described above. We see that the tree pair diagram for the \( c \) part has enough carets so that only right carets need to be added to the trees in the diagrams for \( p \) and \( q \), so that the diagram constructed will indeed have the original word as its \( pcq \) factorization.

We can compute the number of carets of a word \( w \in F \) algebraically from the normal form of \( w \) [4].
Proposition 3.5 ([4]). Given a positive word in \( w \in F \) in the form
\[
w = x_{i_1}^{r_1}x_{i_2}^{r_2} \ldots x_{i_n}^{r_n},
\]
then the number of carets \( N(w) \) in either tree of a reduced tree diagram representing \( w \) is
\[
N(w) = \max\{i_k + r_k + \ldots + r_n + 1, \text{ for } k = 1, 2, \ldots, n\}.
\]

When \( w \in F \) is not a positive word, \( N(w) \) is the maximum of the two numbers obtained by applying Proposition 3.5 to the positive and negative parts of the normal form for \( w \). When considering elements of \( T \), we recall that the number of carets in a tree pair diagram for \( c_j^i \) is equal to \( i + 1 \). Thus it is always possible to decide algebraically when \( w \in T \) written in \( pcq \) form corresponds to a tree pair diagram, using Proposition 3.5 to count the carets for the positive and negative parts of the word.

4. Normal forms in \( T \)

In \( T \), we will declare the words in \( pcq \) form which are \( pcq \) factorizations associated to reduced diagrams to be the normal forms for elements of \( T \), similar to the approach used in \( F \). However, it is no longer true that these words cannot be shortened by applying a relation. As we saw with the normal form \( x_1c_2x_1^{-1} \) in \( T \), a word may be the shortest word representing an element which satisfies the factorization condition, yet there may be shorter words we can obtain by applying a relator which do not satisfy the factorization condition.

Thus, when algebraically characterizing the normal form for elements of \( T \), we restrict ourselves to words of \( pcq \) form which satisfy the factorization condition, regardless of whether or not a relator may reduce the length of the word. We next need to specify algebraic conditions which characterize the \( pcq \) forms that correspond to normal forms.

Theorem 4.1. Let \( w \) be a \( pcq \) factorization for an element \( g \in T \) associated to a marked tree pair diagram in which each tree has \( i + 1 \) carets, where the \( c \) part of the word is \( c_j^i \) with \( 1 \leq j \leq i + 1 \). A reduction of a pair of carets from the tree pair diagram occurs only if the word \( w \) satisfies one of the following conditions:

1. There exists a pair of generators \( x_{k+j} \) and \( x_k^{-1} \), with \( 0 \leq k \leq i - j - 1 \), and neither of the two generators \( x_{k+j+1} \) and \( x_{k+1}^{-1} \) appear. The reduction corresponds to applying the relation
\[
x_{k+j}c_j^ix_k^{-1} = c_{j-1}^i
\]
after applying relations from \( F \) in the \( p \) and \( q \) parts of the word, if necessary, to make \( x_{k+j} \) and \( x_k^{-1} \) adjacent to \( c_j^i \).
(2) The generator $x_k^{-1}$ with $k = i - j$ appears, and $x_{k+1}^{-1}$ does not. The reduction corresponds to applying
\[ c_i^j x_k^{-1} = \frac{c_i^{j-1}}{c_{i-1}} \]
after possibly using relations from $F$ as in (1).

(3) There exists a pair of generators $x_{k-i+j-2}$ and $x_k^{-1}$ for $i > k \geq i - j + 2$ and neither one of the generators $x_{k-i+j-1}$ or $x_{k+1}^{-1}$ appear. The reduction corresponds to applying
\[ x_{k-i+j-2} c_i^j x_k^{-1} = \frac{c_i^{j-1}}{c_{i-1}} \]
after possibly applying relations from $F$.

(4) The generator $x_{j-2}$ appears, and the generator $x_{j-1}$ does not appear. The reduction corresponds to
\[ x_{j-2} c_i^j = \frac{c_i^{j-1}}{c_{i-1}} \]
after possibly applying relations from $F$.

Proof. Let $g \in T$ be represented by a tree pair diagram $(T_-, T_+)$. If both trees have an exposed caret whose leaves are identically numbered, then we call that a reducible caret, as it must be removed in order to obtain the reduced tree pair diagram representing $g$. We now consider algebraic conditions corresponding to a reducible caret in a tree pair diagram.

In the tree pair diagram $(T_-, T_+)$ for $g \in T$, there are two ways of labelling the leaves in the target tree $T_+$. The first labelling corresponds to the order in which the intervals in the subdivisions determined by these trees are paired in the homeomorphism, and is called the cyclic labelling. The cyclic labelling gives the marked leaf in the target tree the number zero, and the other leaves are given increasing labels from left to right around the leaves of the tree. The second labelling ignores the marking and puts the leaves in increasing order from left to right, beginning with zero. The first labelling is used to determine which leaves in $T_-$ are paired with which leaves in $T_+$, and the second labelling is used in the computation of leaf exponents to determine the powers of the generators that appear in the word.

Suppose that the tree pair diagram for $g \in T$ is not reduced. The four cases above correspond to the following four possible locations of a reducible caret relative to the marked leaf in the target tree.

(1) Case (1) corresponds to the case when the left leaf of the reducible caret is to the left of the marked leaf in $T_+$, but the reducible caret is not the rightmost caret in $T_-$.

(2) Case (2) corresponds to the special case when the reducible caret is a right caret in $T_-$ (in which case necessarily the left leaf is to the
left of the marked leaf in \( T_+ \). Leaf exponents from right carets will always be zero and thus right carets cannot contribute generators to the normal form. They may still result in an exposed reducible caret, which occurs exactly in this case, and the reduction will only affect the \( q \) part of the normal form.

(3) Case (3) corresponds to the case when the left leaf of the reducible caret is either to the right of or coincides with the marked leaf in \( T_+ \), but the reducible caret is not the rightmost caret in \( T_+ \).

(4) Case (4) corresponds to the special case when the reducible caret is a right caret in \( T_+ \) (in which case it cannot be to the left of the marked caret in \( T_+ \)). As in Case (2), the exposed caret in this case is a right caret and does not contribute a generator to the normal form, but may still be reduced. This cancellation affects only the \( p \) part of the normal form.

We will prove case (1), and the proofs in the other cases are analogous. We consider an element \( g \in T \) represented by a tree pair diagram. If \( g \notin F \) then the two labellings of the leaves of \( T_+ \) do not coincide. It is easy to see that if \( w \) is the \( pcq \) word satisfying the factorization condition, where the middle expression for \( c \) is \( c_j^l \), then the marked leaf, with leaf number zero in the first labelling, always corresponds to the generator \( x_{i-j+2} \). Thus the leaf in \( T_+ \) which corresponds to the generator \( x_k^{-1} \) is the one numbered \( k+j \) in the cyclic labelling, and hence the exposed caret in \( T_- \) corresponds to the generator \( x_{k+j} \). Since the caret in question is not a right caret in \( T_- \), the generator \( x_{k+j} \) will appear in \( w \). Since the caret in question is to the left of the marked caret in \( T_+ \), it cannot be a right caret in \( T_+ \), and hence the generator \( x_k^{-1} \) appears in \( w \). The fact that the generator \( x_{k+j+1} \) (respectively \( x_{k+1}^{-1} \)) does not appear in \( w \) follows from the fact that the caret is exposed in the source (respectively target) tree. This proves case (1).

We note that in case (2), the generator \( x_k^{-1} \) is the highest index generator with a negative exponent. This \( x_k^{-1} \) generator must correspond to a caret in \( T_+ \) which is immediately before the marked caret, and its corresponding caret in \( T_- \) is the rightmost caret. Since right carets do not correspond to algebraic generators in the normal form, there is no generator in the positive part of the word involved in this reduction.

Finally, we observe one impossible situation for a marked tree pair diagram, which does not appear in the classification above. It is impossible to have a caret in \( T_+ \) corresponding to the generator \( x_{i-j+1}^{-1} \). The leaf numbered \( i-j+1 \) in the left to right labelling of \( T_+ \) is labelled \( i+1 \) in the cyclic labelling, since leaf \( i-j+2 \) is the marked leaf in \( T_+ \), so the corresponding caret cannot be exposed in \( T_- \). \( \square \)
The conditions in Theorem 4.1 together with the factorization condition algebraically characterize our normal forms. The normal forms for elements in $F$ have already been characterized, so we restrict to elements not in $F$ in our description.

**Theorem 4.2.** Any element $g \in T$ which is not an element of $F$ admits an expression of the form $pcq$ where

$$p = x_{i_1}^{r_1}x_{i_2}^{r_2} \cdots x_{i_n}^{r_n}, \quad c = c_1^j, \quad q = x_{j_1}^{-s_1} \cdots x_{j_m}^{-s_m}x_{j_1}^{s_1},$$

$0 \leq i_1 < i_2 < \cdots < i_n$, $0 \leq j_1 < j_2 < \cdots < j_m$, and $1 \leq j < i + 2$. Among all the words in this form representing an element, there is a unique one satisfying the following conditions, and it is the normal form.

- The factorization condition, which we now state as $i + 1 \geq \max\{N(p), N(q)\}$.
- The word does not admit any reductions. Namely, this word satisfies the following conditions:
  - If there exists a pair of generators $x_{k+j}$ and $x_{k}^{-1}$ simultaneously, for $k \leq i - j - 1$, then one of the generators $x_{k+j+1}$ or $x_{k+1}^{-1}$ must appear as well.
  - If there is a generator $x_{k}^{-1}$ with $k = i - j$, then $x_{k+1}^{-1}$ must exist too.
  - If there exists a pair of generators $x_{k-i+j-2}$ and $x_{k}^{-1}$ for $k \geq i - j + 2$, then one of the generators $x_{k-i+j-1}$ or $x_{k+1}^{-1}$ must appear as well.
  - If there exists a generator $x_{j-2}$, then a generator $x_{j-1}$ must also appear.

**Proof.** We claim that the conditions above precisely describe a set of unique normal forms for $T$. A $pcq$ word satisfying the factorization condition is the $pcq$ factorization associated to a marked tree pair diagram. However, if the $pcq$ word satisfies all four reduction conditions, we have just shown in the previous theorem that this diagram is in fact the unique reduced diagram, and hence the word is in fact a normal form. \qed

We remark that the Pumping Lemma together with the reductions in Theorem 4.1 give an explicit way of algebraically transforming any word in the generators of $T$ into a normal form. Namely, given any word, we rewrite it in $pcq$ form using the process described following the Pumping Lemma. If the resulting word does not satisfy the factorization condition, then we iterate the Pumping Lemma until we obtain a word for which the factorization condition is satisfied. The Pumping Lemma increases the number of carets for $c$ and the number of carets for one of the words $p$ and $q$. Once a word is obtained which satisfies the factorization condition,
there must be a corresponding tree pair diagram for the element. Now, if
the word satisfies any of the reduction conditions in Theorem 4.1, we apply
them successively using the relations described there. This method thus
produces the unique normal form.

5. The Word Metric in $T$

5.1. Estimating the Word Metric. For metric questions concerning $T$,
we must consider a finite generating set instead of the one used to obtain the
normal form for elements. We now approximate the word length of an ele-
ment of $T$ with respect to the generating set $\{x_0, x_1, c_1\}$, using information
contained in the normal form and the tree pair diagram. These estimates
are similar to those for the word metric in $F$ with respect to the generating
set $\{x_0, x_1\}$ found in [3] and [4].

**Theorem 5.1.** Let $w \in T$ have normal form
\[ w = x_{i_1}^{r_1} x_{i_2}^{r_2} \ldots x_{i_n}^{r_n} c_j^{i_m} x_{j_1}^{-s_2} \ldots x_{j_m}^{-s_1}. \]

We define
\[ D(w) = \sum_{k=1}^n r_k + \sum_{l=1}^m s_l + i + j + 1. \]

Let $|w|$ denote the word metric in $T$ with respect to the generating set
$\{x_0, x_1, c_1\}$. There exists a constant $C > 0$ so that for every $w \in T$,
\[ \frac{D(w)}{C} \leq |w| \leq C D(w) \]
and similarly, for $N(w)$ the number of carets in a reduced tree pair diagram
representing $w$,
\[ \frac{N(w)}{C} \leq |w| \leq C N(w). \]

**Proof.** These inequalities follow from the correspondence between the nor-
mal form and the tree pair diagram for an element $w \in T$. It is clear, from
Proposition 3.5, that $N(w) \geq \sum_{k=1}^n r_k$, $N(w) \geq \sum_{l=1}^m s_l$, $N(x) \geq i$, and
$N(w) \geq j$. The inequality $N(w) \geq i$ is clear from the fact that $c_1$ has $i + 1$
carets. These inequalities prove that
\[ D(w) \leq 5 N(w). \]

We rewrite the generators $x_i$ and $c_j$ in terms of $x_0$, $x_1$ and $c$ and look at
the lengths of these words to obtain the inequality
\[ |w| \leq C D(w) \]
for some constant $C > 0$. Combining the two inequalities above we have
\[ |w| \leq C' N(w). \]
To obtain lower bound on the word length, one only needs to observe that the tree pair diagram for each generator has either two or three carets. If \( u \) is a word in \( x_0, x_1 \) and \( c \) with length \( n \), then as these generators are multiplied together, each product may add at most 3 carets to the tree pair diagram. Thus the diagram for \( u \) will have at most 3 \( n \) carets. It then follows that

\[ N(w) \leq 3|w|. \]

Combining this with the above inequality, we obtain the desired bounds. \( \square \)

We use Theorem 5.1 to show that the inclusion of \( F \) in \( T \) is a quasi-isometric embedding. This means that there are constants \( K > 0 \) and \( C \) so that for any \( w, z \in F \) we have

\[ \frac{1}{K} d_F(w, z) - C \leq d_T(w, z) \leq K d_F(w, z) + C \]

where \( d_F \) and \( d_T \) represent the word metric in \( F \) and \( T \) respectively, with regard to the generating set \( \{x_0, x_1\} \) of \( F \) and \( \{x_0, x_1, c\} \) of \( T \).

When considering whether the inclusion of a finitely generated subgroup \( H \) into a finitely generated group \( G \) is a quasi-isometric embedding, we can instead equivalently show that the distortion function is bounded. The distortion function is defined by

\[ h(r) = \frac{1}{r} \max \{|x|_H \text{ such that } x \in H, \ |x|_G \leq r \}. \]

Word length in \( F \) is comparable to the number of carets in the reduced tree pair diagram representing the word, as seen in [4, 8]. This, combined with Theorem 5.1 easily shows that the distortion function is bounded, and thus proves the following corollary.

**Corollary 5.2.** The inclusion of \( F \) in \( T \) is a quasi-isometric embedding.

5.2. Comparing word length in \( F \) and \( T \). Although Corollary 5.2 shows that \( F \) is quasi-isometrically embedded in \( T \), we now show that the word length of certain elements of \( F \) does not change when these elements are considered as elements of \( T \), with respect to natural finite generating sets. For this, we use the standard finite generating set \( \{x_0, x_1\} \) for \( F \) and the generating set \( \{x_0, x_1, t\} \) for \( T \), where \( t = c_0 \) is the non-identity element of \( T \) in which each tree has a single caret. This element corresponds to a rotation of the unit circle of order 2. We use \( t \) instead of \( c = c_1 \) for the third generator because we are interested in understanding how multiplication by generators can change the number of carets, which is more straightforward using \( t \) than \( c \).

To find elements which have the same length in \( F \) and \( T \) with respect to these generating sets, we consider the process by which they are constructed.
A geodesic word \(a_n \cdots a_2 a_1\) in the generators \(\{x_0, x_1, t\}\) representing \(w \in F \subset T\) describes how a tree pair diagram for \(w = (T_-, T_+)\) is created by successively applying the generators \(a_k\) to the tree pair diagram for the word \(a_{k-1} \cdots a_2 a_1\), as \(k\) increases from 1 to \(n\). We begin with the identity element of \(T\) and its reducible tree pair diagram consisting of two identical trees with one caret each, and multiply first by \(a_1\). This changes the tree pair diagram of the identity element by creating some new carets or adding a marking. As each successive generator is added to the product, the number of carets in the existing tree pair diagram may increase, decrease or remain the same. If the number of carets increases, it can increase by at most two since each generator has at most two carets in addition to the root caret. In other cases, the number of carets may remain the same or decrease. When all generators in the sequence \(a_n \cdots a_2 a_1\) have been added to the product, the resulting tree pair diagram is \((T_-, T_+)\). We now carefully analyze the circumstances under which a single generator in this product can add two carets to an existing tree pair diagram.

**Lemma 5.3.** If \(w \in T\) is a non-identity element, and \(w \neq t\), then \(N(tw) = N(w)\).

**Proof.** Since the tree pair diagram for \(t\) contains a single caret, no new carets must be added to the tree pair diagram for \(w\) in order to perform the multiplication \(tw\). It is easy to see that no reduction can occur after multiplication by \(t\). \(\square\)

**Lemma 5.4.** Let \(w = (T_-, T_+) \in T\) be nontrivial and \(\alpha \in \{x_0^{\pm 1}, x_1^{\pm 1}, t\}\). Then \(N(\alpha w) = N(w) + 2\) if and only if \(\alpha = x_1^{\pm 1}\) and the right subtree of the root caret of \(T_+\) is empty.

**Proof.** Since the source tree of the tree pair diagram for \(x_1^{\pm 1}\) contains two carets in the right subtree of the root caret, it is clear that these carets must be added to \((T_-, T_+\) in order to perform the multiplication, and that the resulting product does not have reducible carets in the right subtree of the root caret in either tree. The tree pair diagram for \(x_0^{\pm 1}\) contains two carets in each tree, but one is the root caret, so the maximum number of carets that could be added to \((T_-, T_+)\) in order to multiply by \(x_0^{\pm 1}\) is one. Lemma 5.3 completes the proof. \(\square\)

We consider the process of constructing the tree pair diagram for a product \(a_n \cdots a_2 a_1\) where \(a_i \in \{x_0^{\pm 1}, x_1^{\pm 1}, t\}\). Let \(P_i = a_i a_{i-1} \cdots a_2 a_1\) for \(1 \leq i \leq n\). We construct the tree pair diagrams for the successive products \(P_i\) for \(i = 1, \cdots, n\). Each additional generator may either reduce, leave unchanged, increase by one, or increase by two the number of carets in the tree pair diagram corresponding to the suffix \(P_i\) of \(P_n\). To distinguish those
generators which add two carets to the tree pair diagram, we will use the letter \(b\), while for other generators we will use the letter \(a\). So we represent an element \(w \in T\) by a string of generators

\[ a_r \cdots a_h b_k \cdots a_{s+1} b_1 a_s \cdots a_1 \]

where if \(a_t\) is the generator immediately to the right of \(b_j\), then

\[ N(b_j a_t \cdots a_1) = N(a_t \cdots a_1) + 2, \]

and otherwise we have

\[ N(a_s \cdots a_1) \leq N(a_{s-1} \cdots a_1) + 1. \]

**Lemma 5.5.** If \(w \in F\) and \(w = a_r \cdots a_h b_k \cdots a_{s+1} b_1 a_s \cdots a_1\) where not all \(a_i \in \{x_0^{\pm 1}, x_1^{\pm 1}\}\), then there must be at least two indices \(i\) and \(j\) so that \(a_i = a_j = t\).

**Proof.** Suppose that in the expression above for \(w \in F\), there was a single letter \(t\). Then we easily obtain \(t \in F\), a contradiction. \(\square\)

**Lemma 5.6.** Let \(w \in T\) be given by an expression of the form \(a_r \cdots a_h b_k \cdots a_{s+1} b_1 a_s \cdots a_1\). Then \(b_{j+1}\) and \(b_j\) cannot be adjacent in the expression. Note that possibly \(s = 0\) in which case the word ends with \(b_1\), and that possibly \(h - 1 = r\) in which case the word begins with \(b_k\).

**Proof.** Let \(v\) be the suffix of the word \(w\) to the right of the generator \(b_j\). We know from Lemma 5.4 that in a tree pair diagram \((R_-, R_+)\) for \(v\), the right subtree of the root caret of \(R_+\) is empty. Since \(b_j = x_1^{\pm 1}\), the right subtree of the root caret of \(S_1\) has one of two forms: two right carets or a right caret with a left child. In either case, the right subtree of the root caret is nonempty, and thus the next generator in the multiplication cannot add two carets to the tree pair diagram, so is not \(b_{j+1}\). \(\square\)

We now characterize one type of element of \(F\) whose word length is unchanged when viewed as an element of \(T\), using the generating set \(\{x_0, x_1\}\) for \(F\) and \(\{x_0, x_1, t\}\) for \(T\). These are elements \(w \in F\) for which \(N(w)\) exceeds the word length \(|w|_F\). Fordham [8] computes \(|w|_F\) by assigning an integer weight between zero and four to each pair of carets in the tree pair diagram representing \(w\). In a given word there are at most two weights of zero. Here we investigate words in which most weights are one. Such words, for example, are represented by tree pair diagrams with no interior carets having right children.

**Theorem 5.7.** If \(w \in F\) with \(N(w) \geq |w|_F + 1\) then \(|w|_T = |w|_F\), where word length if computed with respect to the generating set \(\{x_0, x_1\}\) for \(F\) and \(\{x_0, x_1, t\}\) for \(T\).

We immediately obtain the following corollary, since \(|x_0^n|_F = |x_1^n|_F = n\), while \(N(x_0^n) = n + 1\) and \(N(x_1^n) = n + 3\).
Corollary 5.8. The elements $x_0^n$ and $x_1^n$ have word length $n$ in both $F$ and $T$ with respect to the finite generating sets $\{x_0, x_1\}$ and $\{x_0, x_1, t\}$ respectively.

Proof. Suppose $w \in F$ can be written as $w = a_r \cdots a_2 a_1$ where $a_i \in \{x_0^{\pm 1}, x_1^{\pm 1}, t\}$, and $r < n = |w|_F$. If $a_i$ is a generator which adds two carets to the tree pair diagram, then rename this generator $b_j$. So we rewrite the expression for $w$ as

$$w = a_p \cdots b_k \cdots a_{s+1} b_1 a_s \cdots a_1$$

with $r = p + k$ where $k$ is the number of $b_j$ generators. By Lemma 5.6 we have $k - 1 \leq p$, and we know from Lemma 5.4 that $b_i = x_i^{\pm 1}$.

We first prove the following lemma relating to this expression for $w$:

Lemma 5.9. Let $a_p \cdots a_k b_k \cdots b_j a_i \cdots a_{s+1} b_1 a_s \cdots a_1$ be an expression for a word $w \in T$, where each $b_j = x_i^{\pm 1}$ is a generator which adds two carets to the tree pair diagram, and each $a_i$ is a generator from the set $\{x_0^{\pm 1}, x_1^{\pm 1}, t\}$ which adds at most one caret to the tree pair diagram. We consider the generators $a_k$ which appear between $b_{j+1}$ and $b_j$.

1. There is at least one generator between $b_{j+1}$ and $b_j$ which does not increase the number of carets.

2. Let $a_{j_1}$ be the generator between $b_{j+1}$ and $b_j$ closest to $b_{j+1}$ such that if $v_{j_1}$ is the suffix of $w$ immediately to the right of $a_{j_1}$, then the right subtree of the target tree of $v_{j_1}$ is nonempty, but the right subtree of the target tree of $a_{j_1} v_{j_1}$ is empty.

If $a_{j_1} = t$ and $b_j$ is not the rightmost generator in the expression for $w$, then there is another generator $a_{j_1}$ between $b_{j+1}$ and $b_j$ which does not add carets to the tree pair diagram.

Proof. We know that $b_{j+1} = x_1^{\pm 1}$, and for multiplication by $b_{j+1}$ to add two carets to the tree pair diagram, the target tree of the tree pair diagram for the suffix of $w$ immediately following $b_{j+1}$ must have an empty right subtree. We also know that the target tree of the tree pair diagram for the suffix of $w$ beginning with $b_j$ does not have an empty right subtree, according to Lemma 5.4. Thus the generator $a_{j_1}$ in the statement of the lemma must exist.

As in the statement of the lemma, let $v_{j_1}$ denote the suffix of $w$ immediately to the right of $a_{j_1}$. We now claim that $N(v_{j_1}) \geq N(a_{j_1} v_{j_1})$. If $a_{j_1} = t$ then the claim follows from Lemma 5.3. If $a_{j_1} = x_0^{-1}$ or $a_{j_1} = x_1^{\pm 1}$, and $N(a_{j_1} v_{j_1}) = N(v_{j_1}) + 1$, then the right subtree of the target tree of $a_{j_1} v_{j_1}$ can never be empty. Thus if $a_{j_1}$ is one of the above three generators, it can never add a caret to the tree pair diagram of $v_{j_1}$, because we know that the right subtree of the target tree of $a_{j_1} v_{j_1}$ must be empty.
Suppose that \( a_J = x_0 \). If \( N(a_J v_J) = N(v_J) + 1 \), then the right subtree of the target tree of the tree pair diagram for \( v_J \) must be empty. But this contradicts the definition of \( a_J \), and thus \( N(a_J v_J) \leq N(v_J) \), and the claim is again true.

Now consider the case \( a_J = t \) when \( b_j \) is not the rightmost generator in the expression for \( w \). Let \( b_j \cdots a_1 = (R_-, R_+) \) and \( v_J = (S_-, S_+) \). Since \( b_j = x_1^{\pm 1} \), the right subtree of the root caret of \( R_+ \) has one of two forms: two right carets or a right caret with a left child. Let \( A \) denote the left subtree of the root caret of this tree. Denote the right subtree of \( S_+ \) by \( A' \). Since \( b_j \) is not the rightmost generator in the expression for \( w \), we know that both \( A \) and \( A' \) are nonempty. Thus there must be at least one generator between \( a_J \) and \( b_j \). We claim that there must be a generator between \( a_J \) and \( b_j \) which does not increase the number of carets in the tree pair diagram.

Since the right subtree of the root caret of \( R_+ \) is nonempty, as is the left subtree \( A \) of the root caret, we see immediately that \( N(x_1^{\pm 1} b_j \cdots a_1) \leq N(b_j \cdots a_1) \). Similarly, multiplication by \( b_j^{-1} \) cannot increase the number of carets in the tree pair diagram \((R_-, R_+)\), and multiplication by \( t \) never increases the number of carets in any tree pair diagram. Thus the only generator which can precede \( b_j \) and increase the number of carets in the tree pair diagram is \( b_j \) itself. Repeat occurrences of \( b_j = x_1^{\pm 1} \) are written as \( a_i \) in the expression for \( w \), since they do not increase the number of carets by 2. But the target tree in the tree pair diagram for \( b_j^m \cdots a_1 \) will always have left subtree \( A \) and a nonempty right subtree. Since \( v_J = (S_-, S_+) \) has a tree \( S_+ \) with empty right subtree of the root caret, we see that there must be another generator not equal to \( b_j \) between \( b_j^m \cdots a_1 \) and the leftmost letter in \( v_J \). By the above argument, this generator must not increase the number of carets in the tree pair diagram. Hence there must be another generator between \( a_J \) and \( b_j \) which does not increase the number of carets in the tree pair diagram. \( \square \)

According to Lemma 5.9, between each pair of generators \( b_{j+1} \) and \( b_j \), there is a generator which does not add carets to the tree pair diagram. There are at least \( k - 1 \) such generators in the expression for \( w \), between \( b_{j+1} \) and \( b_j \) for \( j = 1, 2, \cdots k - 1 \). In the proof of Lemma 5.9, these generators are denoted \( a_J \). Lemma 5.9 also shows the existence of at least one additional generator which does not increase the number of carets in the tree pair diagram, for the following reason.

We can assume that the expression for \( w \) contains at least one generator \( t \), otherwise we would have a word in \( F \). In \( F \), we know that the word length of \( w \) is \( n > r \). It follows from Lemma 5.5 that there must be at least two \( t \) generators in the expression for \( w \). We now prove in
three cases that there must be an additional generator in the sequence
\(a_p \cdots a_h b_k \cdots b_j a_t \cdots a_{s+1} b_1 a_s \cdots a_1\) which is not one of the \(a_J\) generators
guaranteed by Lemma 5.9, which does not increase the number of carets in
the tree pair diagram.

(1) If no \(a_J = t\), then we have found two additional generators which do
not increase the number of carets in the tree pair diagram, namely
the two instances of the generator \(t\) guaranteed by Lemma 5.5.

(2) If two \(t\) generators play the role of \(a_J\), one between \(b_{j+1}\) and \(b_j\),
and the other between \(b_{m+1}\) and \(b_m\), then at most one of \(b_j\) and \(b_m\),
say \(b_j\), can be the rightmost generator in the sequence representing
\(w\). Then Lemma 5.9 guarantees one additional generator between
\(b_{m+1}\) and \(b_m\) which does not increase the number of carets in the
tree pair diagram.

(3) If only one of the \(t\) generators in the expression is an \(a_J\) genera-
tor, then the other \(t\) generator guaranteed by Lemma 5.5 does not
increase the number of carets in the tree pair diagram.

Thus we always have at least \(k\) generators in the expression
\(a_p \cdots a_h b_k \cdots b_j a_t \cdots a_{s+1} b_1 a_s \cdots a_1\) which do not increase the number of carets in the tree pair diagram.

We then obtain the following upper bound for the number of carets \(N(w)\):
\[
n + 1 \leq N(w) \leq 1 + 2k + (r - k - k) = r + 1 < n + 1
\]
which is clearly a contradiction. Thus we are unable to express \(w\) using
fewer generators than \(n = |w|_F\).

5.3. Cyclic subgroups of \(T\). We can use the interpretation of elements
of \(T\) as homeomorphisms of the circle to understand how \(F\) is contained in
\(T\) and that the cyclic subgroups of \(T\) are quasi-isometrically embedded.

**Theorem 5.10.** Every element in \(T\) has a power which is conjugate to an
element in \(F\).

**Proof.** Ghys and Sergiescu [9] and Liousse [12] show that every element in \(T\)
has rational rotation number as a homeomorphism of \(S^1\). A standard result
in dynamics is that a homeomorphism of the circle has rational rotation
number if and only if it has a periodic orbit (see, for example Katok and
Hasselblatt [11].) Thus, every element in \(T\) has a periodic orbit. That is,
given \(f \in T\), there exists a dyadic rational \(x \in [0, 1]\) such that \(f^n(x) = x\)
for some \(n\).

Given a dyadic rational \(x\), we consider the rotation of the circle \(t_x(p) =
p + x\), where we view \(S^1\) as the unit interval with the endpoints identified.
Then $t_x$ is an element of $T$, and conjugation of $w \in T$ by $t_x$, for the appropriate choice of $x$ yields an element with zero as a fixed point. Thus the conjugated element must lie in $F$. □

We use Theorem 5.10 to show that cyclic subgroups of $T$ are quasi-isometrically embedded. We again bound the distortion function, as in the proof of Corollary 5.2.

**Theorem 5.11.** Let $x \in T$ be a non-torsion element. Then the cyclic subgroup $\langle x \rangle$ generated by $x$ is quasi-isometrically embedded in $T$.

**Proof.** If the element $x$ lies in $F$, the result can be deduced from Corollary 5.2, since Burillo [3] shows that all cyclic subgroups in $F$ are quasi-isometrically embedded, and $F$ is quasi-isometrically embedded in $T$. If $x \notin F$, it has a power $w^n$ which is conjugate to an element of $F$ by some element $z \in T$. Since $x$ is not a torsion element, this power is not conjugate to the identity element of $F$.

The cyclic subgroup generated by $zw^n z^{-1}$ is quasi-isometrically embedded in $T$. Conjugation by $z$ changes the number of carets by at most a factor of $2N(z)$, and thus we see that $\langle x \rangle$ must also be quasi-isometrically embedded in $T$. □

### 6. Torsion elements

Although the group $F$ is torsion free, both $T$ and $V$ contain torsion elements. It is easy to construct torsion elements in $T$ or $V$ by choosing any binary tree $S$ and making any marked tree pair diagram with $S$ as both source and target tree. If the labelling of the target tree is the same as the labelling of the source tree, we get an unreduced representative of the identity; otherwise, we get a non-trivial torsion element. If this is an element of $T$, the tree pair diagram has $pcq$ factorization in which $q = p^{-1}$. In fact, any torsion element can be represented by such a tree pair diagram, though its reduced marked tree pair diagram may well not have the same source and target trees, corresponding to the fact that although it has a $pcq$ word where $q = p^{-1}$, the normal form may well not have this special balanced appearance.

**Proposition 6.1.** If $f \in F, T$ or $V$ is a torsion element, then it can be represented by a (marked) tree pair diagram with the same source and target trees.

Before proving Proposition 6.1, we establish some notation which links the analytic and geometric interpretations of these groups. For $f \in F, T$, or $V$, if $(T_-, T_+)$ is a marked tree pair diagram representing $f$, then it is sometimes convenient to denote the tree $T_+$ by $f(T_-)$. The element $f$ can
be thought of as mapping the leaves of \( T_+ \) to the leaves of \( f(T_-) = T_+ \), where the marking defines this mapping of the leaves.

Given two rooted binary trees \( T \) and \( T' \), we say that \( T' \) is an expansion of \( T \) if \( T' \) can be obtained from \( T \) by attaching the roots of additional trees to some subset of the leaves of \( T \). We observe that if \((T, f(T))\) is a marked tree pair diagram for \( f \), and \( T'' \) is an expansion of \( T \), then there is always a tree pair diagram \((T', f(T'))\) for \( f \), and \( f(T') \) is an expansion of \( f(T) \). Given two rooted binary trees \( S \) and \( T \), by the minimal common expansion of \( S \) and \( T \) we mean the smallest rooted binary tree which is an expansion of both \( S \) and \( T \). Using this language, if \((T, f(T))\) and \((S, g(S))\) are marked tree pair diagrams for \( f \) and \( g \) respectively, the process described in Section 2.3 for creating a tree pair diagram for the product \( gf \) could be summarized as follows. If \( E \) is the minimal common expansion of \( f(T) \) and \( S \), then there are tree pair diagrams \((f^{-1}(E), E)\) for \( f \), \((E, g(E))\) for \( g \), and \((f^{-1}(E), g(E))\) for \( gf \).

**Proof.** Suppose that \( f \) is a torsion element, and that \( f \) cannot be represented by a tree pair diagram with the same source and target trees. Since there exists a positive integer \( m \) such that \( f^m \) is the identity, it follows that all tree pair diagrams for \( f^m \) have the same source and target trees. We reach a contradiction by constructing (marked) tree pair diagrams \((A_n, B_n)\) for \( f^n \) such that \( A_n \neq B_n \) for every \( n \geq 1 \). These tree pair diagrams are constructed inductively, viewing \( f^n \) as a product \((f^{n-1})(f)\). For \( n = 1 \), let \((A_1, B_1)\) be the reduced marked tree pair diagram for \( f \). If \( k \geq 2 \), suppose the marked tree pair diagram \((A_{k-1}, B_{k-1})\) for \( f^{k-1} \) has been constructed.

Let \( E_{k-1} \) be the minimal common expansion of the trees \( A_1 \) and \( B_{k-1} \). Then \( f^k \) has tree pair diagram \((f^{-k+1}(E_{k-1}), f(E_{k-1}))\), and we let \( B_k = f(E_{k-1}) \) and \( A_k = f^{-k+1}(E_{k-1}) \).

By construction, \( A_{k+1} \) is an expansion of \( A_k \) for all \( k \geq 1 \). We claim also that \( B_{k+1} \) is an expansion of \( B_k \) for all \( k \geq 1 \). For \( k = 1 \), \( E_1 \) is by definition an expansion of \( A_1 \), which implies that \( B_2 = f(E_1) \) is an expansion of \( B_1 = f(A_1) \). Suppose inductively that \( B_k \) is an expansion of \( B_{k-1} \). Now \( E_k \) is an expansion of \( B_k \) and \( A_1 \), so \( E_k \) is an expansion of \( B_{k-1} \) and \( A_1 \). But \( E_{k-1} \) is the minimal common expansion of \( B_{k-1} \) and \( A_1 \), so \( E_k \) is an expansion of \( E_{k-1} \), which implies that \( B_{k+1} = f(E_k) \) is an expansion of \( B_k = f(E_{k-1}) \).

Now if \( A_n = B_n \), then since \( A_n \) is an expansion of \( A_1 \), \( B_n \) is an expansion of \( A_1 \). But since \( E_{n-1} \) is the minimal common expansion of \( B_{n-1} \) and \( A_1 \), this implies that \( B_n = f(E_{n-1}) \) is an expansion of \( E_{n-1} \). But they have the same number of carets, so in fact \( f(E_n) = E_{n-1} \), which cannot be. Hence \( A_n \neq B_n \) for all \( n \), as claimed. \( \square \)
Corollary 6.2. An element of $T$ is torsion if and only if it is a conjugate of some $c^j_i$.

Proof. If an element is torsion, then it admits a diagram with two equal trees. The $pcq$ factorization associated with this diagram has the form $pcq^{-1}$, where $p$ is a positive element of $F$. \hfill $\Box$

A particularly natural torsion subgroup is the subgroup $R$ of pure rotations, where by a pure rotation by $d = \frac{a}{2n}$ (where $a$ is not divisible by 2) we mean an element $g_d \in T$ which corresponds to the homeomorphism of $S^1$ given by

$$g_d(t) = \begin{cases} t + \frac{a}{2n} & \text{if } 0 \leq t < 1 - \frac{a}{2n} \\ t + \frac{a}{2n} - 1 & \text{if } 1 - \frac{a}{2n} \leq t < 1 \end{cases}$$

Such pure rotations were used in Section 5.3 to conjugate the fixed point of a homeomorphism to 0.

This subgroup is isomorphic to the group of dyadic rational numbers modulo 1, which has a 2-adic metric as follows: if $x = \frac{p}{2^k}$, $y = \frac{q}{2^m}$, and $z = |x - y| = \frac{r}{2^n}$, where $p, q$ and $r$ are odd, then $d(x, y) = 2^k$. With respect to this metric, the subgroup of rotations is quasi-isometrically embedded in $T$.

Proposition 6.3. The subgroup $R$ of the pure rotations, with the 2-adic metric, is quasi-isometrically embedded in $T$.

Proof. We note that if $g \in T$ is the rotation by $\frac{a}{2n}$ where $a$ is not divisible by 2, then there are $2^n - 1$ carets in the reduced tree pair diagram representing $g$, so $N(g) = 2^n - 1$. Since we have shown that the word length of $g$ in $T$ is bi-Lipschitz equivalent to $N(g)$, the proposition follows. \hfill $\Box$

References


Dep. de Matemàtica Aplicada IV, Univ. Politècnica de Catalunya, Escola Politècnica Superior de Castelldefels, 08860 Castelldefels, Barcelona, Spain
E-mail address: burillo@mat.upc.es

Department of Mathematics, The City College of New York & The CUNY Graduate Center, New York, NY 10031
E-mail address: cleary@sci.ccny.cuny.edu

Department of Mathematics, Trinity College, Hartford, CT 06106
E-mail address: melanie.stein@trincoll.edu

Department of Mathematics, Bowdoin College, Brunswick, ME 04011
E-mail address: jtaback@bowdoin.edu