A FINITELY PRESENTED GROUP OF PIECEWISE PROJECTIVE HOMEOMORPHISMS

YASH LODHA AND JUSTIN TATCH MOORE

This paper is dedicated to the memory of William Thurston (1946–2012).

Abstract. In this article we will describe a finitely presented subgroup of Monod’s group of piecewise projective homeomorphisms of \( \mathbb{R} \). This in particular provides a new example of a finitely presented group which is nonamenable and yet does not contain a nonabelian free subgroup. It is in fact the first such example which is torsion free. We will also develop a means for representing the elements of the group by labeled tree diagrams in a manner which closely parallels Richard Thompson’s group \( F \).

1. Introduction

The notion of an amenable group was introduced by von Neumann as an abstract means for preventing the existence of paradoxical decompositions of the group: a discrete group is amenable if it admits a finitely additive translation invariant probability measure. At the heart of Banach and Tarski’s paradoxical decomposition of the sphere is the existence of a paradoxical decomposition of the free group on two generators. Since subgroups of amenable groups are easily seen to be amenable, it is natural to ask whether every nonamenable group contains a free group on two generators. Day was the first to pose this problem in print [8], where he attributed it to John von Neumann.

In 1980, Ol’shanskii solved the von Neumann-Day problem by producing a counterexample [13]. Soon after, Adyan showed that certain Burnside groups are also counterexamples [1][2]. These examples are not finitely presented and the restriction of the von Neumann-Day problem to the class of finitely presented groups remained open until Ol’shanskii and Sapir constructed an example in 2003 [14]. Shortly after, Ivanov published another finitely presented counterexample [10], that is somewhat simpler but similar in spirit to the Ol’shanskii-Sapir example. Both examples were produced by elaborate inductive constructions and are difficult to analyze. It is also interesting to note that both of these examples are based on the constructions of nonamenable torsion groups (they are torsion-by-cyclic) and in particular are far from being torsion free.

In his recent article [12], Monod produced a new family of counterexamples to the von Neumann-Day problem. They are all subgroups of the group \( H \) consisting

2010 Mathematics Subject Classification. Primary: 43A07; Secondary: 20F05.

Key words and phrases. amenable, finitely presented, free group, piecewise, projective, Thompson’s group, torsion free.

This research was supported in part by NSF grant DMS–1262019. The second author would like to thank Alain Valette for a number of helpful conversations concerning Monod’s example in [12] which took place while the author was visiting Université de Neuchâtel in January of 2013.
of all piecewise projective transformations of the real projective line which fix the point at infinity. Monod demonstrated that \( H \) does not contain nonabelian free subgroups by adapting the method of Brin and Squier [4].

In this article, we will isolate a finitely presented nonamenable subgroup of Monod’s group \( H \). To our knowledge, this provides the first finitely presented torsion free counterexample to the von Neumann-Day problem. Moreover, our presentations for this group are very explicit: it has a presentation with 3 generators and 9 relations as well as a natural infinite presentation. The group is generated by \( a(t) = t + 1 \) together with the following two homeomorphisms of \( \mathbb{R} \):

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

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

The main result of this paper is the following.

**Theorem 1.1.** The group \( G_0 \) generated by the functions \( a(t), b(t), \) and \( c(t) \) is nonamenable and finitely presented.

Since it is a subgroup of \( H \), \( G_0 \) does not contain a nonabelian free subgroup. We claim no originality in our proof that \( G_0 \) is nonamenable; this is a routine modification of the methods of [12] which in turn relies on [6] [7].

It is interesting to note that, by an unpublished result of Thurston, \( a(t) \) and \( b(t) \) generate the subgroup \( P(\mathbb{Z}) \leq H \), consisting of those homeomorphisms which are \( C^1 \) and piecewise \( \text{PSL}_2(\mathbb{Z}) \), which he moreover showed is a copy of Richard Thompson’s group \( F \). In fact the methods of [12] easily show that \( t \mapsto t + 1/2 \) and \( b(t) \) generate a nonamenable group, although at present it is unclear whether this group is finitely presented. While there are strong parallels between the group \( \langle a, b, c \rangle \) and \( F \), neither [12] nor the present article seems to shed any light on whether \( F \) is amenable.

The paper is organized as follows. Section 2 contains a review of some of the preliminaries we will need later in the paper. Both an infinite and a finite presentation are described in Section 3 and it is demonstrated there how to prove that the finite presentation generates the infinite presentation. Tree diagrams for elements of the group are developed in Section 4. Finally, Section 5 contains a proof that the relations isolated in Section 3 suffice to give a presentation for \( \langle a, b, c \rangle \).

2. Preliminaries

Our analysis of the group \( G_0 \) will closely parallel the now well-established analysis of Thompson’s group \( F \). We direct the reader to the standard reference [3] for the properties of \( F \); additional information can be found in [5]. We shall mostly follow the notation and conventions of [3] [4].

We will take \( \mathbb{N} \) to include 0; in particular all counting will start at 0. Let \( 2^\mathbb{N} \) denote the collection of all infinite binary sequences and let \( 2^{<\mathbb{N}} \) denote the collection of all finite binary sequences. If \( i \in \mathbb{N} \) and \( u \) is a binary sequence of length at least \( i \), we will let \( u \restriction i \) denote the initial part of \( u \) of length \( i \). If \( s \) and \( t \) are finite binary sequences, then we will write \( s \subseteq t \) if \( s \) is an initial segment of \( t \) and \( s \subset t \) if \( s \) is a proper initial segment of \( t \). If neither \( s \subseteq t \) nor \( t \subseteq s \), then we will say that \( s \) and \( t \) are incompatible. The set \( 2^{<\mathbb{N}} \) is equipped with a lexicographic
order defined by \( s \leq_{\text{lex}} t \) if \( t \subseteq s \) or \( s \) and \( t \) are incompatible and \( s(i) < t(i) \) where \( i \) is minimal such that \( s(i) \neq t(i) \).

If \( s \) and \( t \) are two sequences (\( s \) is finite but \( t \) may be infinite), then \( s \sim t \) will be used to denote the concatenation of \( s \) and \( t \). In some circumstances, \( \sim \) will be omitted; for instance we will often write \( s01 \) instead of \( s \sim 01 \). If \( \xi \) and \( \eta \) are infinite sequences, then we will say that \( \xi \) and \( \eta \) are \textit{tail equivalent} if there are \( s, t, \) and \( \zeta \) such that \( \xi = s \sim \zeta \) and \( \eta = t \sim \zeta \). Given an infinite sequence \( s, \bar{s} \) is the sequence satisfying \( \bar{s}(i) = 0 \) if \( s(i) = 1 \) and \( \bar{s}(i) = 1 \) if \( s(i) = 0 \). The constant sequences \( 000..., 111... \) are denoted by \( \bar{0}, \bar{1} \) respectively.

Let \( \mathcal{T} \) denote the collection of all finite rooted ordered binary trees. More concretely, we view elements \( T \) of \( \mathcal{T} \) as \textit{prefix} sets — those sets \( T \) of finite binary sequences with the property that every infinite binary sequence has a unique initial segment in \( T \). Observe that, for each \( m \), there are only finitely many elements of \( \mathcal{T} \) with \( m \) elements. There is also a natural ordering on \( \mathcal{T} \), which we will refer to as \textit{dominance}: if every element of \( S \) has an extension in \( T \), then we say that \( S \) is dominated by \( T \). Notice that if \( S \) is dominated by \( T \), then \(|S| \leq |T| \). If \( A \) is a finite set of binary sequences, then there is a unique minimal element \( T \) of \( \mathcal{T} \) (with respect to the order of dominance) such that every element of \( A \) has an extension in \( T \).

A \textit{tree diagram} is a pair \((L, R)\) of elements of \( \mathcal{T} \) with the property that \(|L| = |R| \). A tree diagram describes a map of infinite binary sequences as follows:

\[
 s_i \sim \xi \mapsto t_i \sim \xi
\]

where \( s_i \) and \( t_i \) are the \( i \)-th elements of \( L \) and \( R \) respectively and \( \xi \) is any binary sequence. The collection of all such functions from \( 2^N \) to \( 2^N \) defined in this way is \textit{Thompson’s group} \( F \). This map is also defined on any finite binary sequence \( u \) such that \( u \) has a prefix in \( L \). We will follow \cite{4} and write \( s.f \) for the result of applying an automorphism \( f \) to the input \( s \). The operation of \( F \) is therefore defined as \( s.(f.g) = (s.f).g \). Thompson’s group \( F \) is generated by the following functions.

\[
 \xi.a = \begin{cases} 
 0\eta & \text{if } \xi = 00\eta \\
 10\eta & \text{if } \xi = 01\eta \\
 11\eta & \text{if } \xi = 1\eta 
\end{cases} \quad \xi.b = \begin{cases} 
 0\eta & \text{if } \xi = 0\eta \\
 10\eta & \text{if } \xi = 10\eta \\
 110\eta & \text{if } \xi = 101\eta \\
 111\eta & \text{if } \xi = 11\eta 
\end{cases}
\]

Recall that the \textit{real projective line} is the set of all lines in \( \mathbb{R}^2 \) which pass through the origin. Such lines can naturally be identified with elements of \( \mathbb{R} \cup \{ \infty \} \) via the \( x \)-coordinate of their intersection with the line \( y = 1 \). In this article it will be useful to represent points on the real projective line by binary sequences derived from their continued fractions expansion. Define a map \( \Phi : 2^N \to \mathbb{R} \cup \{ \infty \} \) as follows. First define \( \phi : 2^N \to [0, \infty] \) by

\[
 \phi(0\xi) = \frac{1}{1 + \frac{1}{\phi(\xi)}} \quad \phi(1\xi) = 1 + \phi(\xi)
\]

and set

\[
 \Phi(0\xi) = -\phi(\xi) \quad \Phi(1\xi) = \phi(\xi).
\]

This function is one-to-one except at \( \xi \) which are eventually constant (i.e. \( r_k = \infty \) for some \( k \)). On sequences which are eventually constant, the map is two-to-one: \( \Phi(s01) = \Phi(s1\bar{0}) \) and \( \Phi(\bar{0}) = \Phi(\bar{1}) = \infty \).
In the mid 1970s, Thurston observed that the functions \( a \) and \( b \) from the introduction become the generators \( a \) and \( b \) for Thompson’s group \( F \) defined above when “conjugated” by \( \Phi \). Moreover, the elements of \( F \) correspond exactly to those homeomorphisms \( f \) of \( \mathbb{R} \) which are piecewise \( \text{PSL}_2(\mathbb{Z}) \) and which have continuous derivatives. We will generally take the viewpoint that \( \Phi \) provides just another way of describing the real projective line, just as decimal expansions allow us to describe real numbers. In particular, we will regard the definitions of \( \text{PSL}_2(\mathbb{Z}) \) and the definitions given above in terms of sequences as being two ways of describing the \textit{same} functions.

Since we will be proving that a group is finitely presented, it will be necessary to deal with formal words over formal alphabets. If \( G \) is a group and \( A \) is a subset of \( G \), an \( A \)-word is a finite sequence of elements of the set \( A \times (\mathbb{Z} \setminus \{0\}) \). We typically denote a pair \((a, n)\) as \( a^n \), but we emphasize here that it is formally distinct from the group element \( a^n \). The \textit{word length} of an \( A \)-word is the sum of the absolute values of the exponents which occur in it.

In order to prove the nonamenability of \( G_0 \), we will need to employ Zimmer’s theory of amenable equivalence relations. Let \( X \) be a Polish space and let \( E \subseteq X^2 \) be an equivalence relation which is Borel and which has countable equivalence classes. \( E \) is \( \mu \)-amenable if, after discarding a \( \mu \)-measure 0 set, \( E \) is the orbit equivalence relation of an action of \( \mathbb{Z} \). (This is not the standard definition, but it is equivalent by [7].) We will need the following two results.

\textbf{Theorem 2.1.} [18] If \( \Gamma \) is a countable amenable group acting by Borel automorphisms on a Polish space \( X \) and \( \mu \) is any \( \sigma \)-finite Borel measure on \( X \), then the orbit equivalence relation is \( \mu \)-amenable.

\textbf{Theorem 2.2.} [9] (see also the discussion in [12]) If \( \Gamma \) is a countable dense subgroup of \( \text{PSL}_2(\mathbb{R}) \), then the action of \( \Gamma \) on the real projective line induces an orbit equivalence relation which is not amenable with respect to Lebesgue measure.

We refer the reader to [11] for further information on amenable equivalence relations.

We will conclude this section by sketching a proof that the group \( G_0 \) from the introduction is nonamenable. Let \( K \) denote the subgroup of \( \text{PSL}_2(\mathbb{R}) \) generated by the matrices

\[
\begin{pmatrix}
1 & 1 \\
0 & 1
\end{pmatrix},
\begin{pmatrix}
0 & 1 \\
-1 & 0
\end{pmatrix},
\begin{pmatrix}
\sqrt{2} & 0 \\
0 & \frac{1}{\sqrt{2}}
\end{pmatrix}.
\]

Viewed as fractional linear transformations, \( K \) is generated by \( t + 1 \), \( 2t \), and \( -1/t \). Since \( K \) contains \( \text{PSL}_2(\mathbb{Z}) \) as a proper subgroup, it is dense in \( \text{PSL}_2(\mathbb{R}) \) and hence by Theorem 2.2, the orbit equivalence relation of its action on the real projective line is not amenable with respect to Lebesgue measure. By Theorem 2.1 it is sufficient to show that \( G_0 \) induces the same orbit equivalence relation on \( \mathbb{R} \setminus \mathbb{Q} \). To see this, it can be verified that the element \( bca^{-1}c^{-1}a \) coincides with \( t \mapsto 2t \) on the interval \([0, 1]\). From this and the identity \( 2(r-n)+2n = 2r \) it follows that the orbits of \( G_0 \) include the orbits of the action of \( \{t \mapsto t+1, \ t \mapsto 2t\} \) on \( \mathbb{R} \setminus \mathbb{Q} \). Finally \( aba \) and \( ba^{-3} \) coincide with \( t \mapsto -1/t \) on the intervals \([-1, -1/2] \) and \([1/2, 1]\) respectively. The assertion about orbits now follows from the fact that for any \( r \in \mathbb{R} \), there is a \( n \in \mathbb{Z} \) such that \( 2^n r \) is in \([-1, -1/2] \cup [1/2, 1] \) and \( (2^n)(-1/(2^n r)) = -1/r \).
3. The presentations

In this section we will describe both a finite and infinite presentation of the group $G_0$ defined in the introduction. We start with the following two primitive functions defined on binary sequences:

$$
\phi(x) = \begin{cases}
0 & \text{if } x = 00 \\
10 & \text{if } x = 01 \\
11 & \text{if } x = 11
\end{cases}
$$

(The function $x$ is nothing but the function $a$ described in the previous section.)

From these functions, we define families of functions $x_s$ ($s \in 2^{<\mathbb{N}}$) and $y_s$ ($s \in 2^{<\mathbb{N}}$) which act just as $x$ and $y$, but localized to those binary sequences which extend $s$.

$$
\xi.x_s = \begin{cases}
\xi.x & \text{if } s = \xi.x \\
\xi & \text{otherwise}
\end{cases}
$$

$$
\xi.y_s = \begin{cases}
\xi.y & \text{if } s = \xi.y \\
\xi & \text{otherwise}
\end{cases}
$$

If $s$ is the empty-string, it will be omitted as a subscript. The relationship between these functions and the functions $a$, $b$, and $c$ of the introduction is expressed by the following proposition.

**Proposition 3.1.** For all $\xi$ in $2^\mathbb{N}$, $\phi(\xi.y) = 2\phi(\xi)$ and

$$
\Phi(\xi).a = \Phi(\xi.x) \quad \Phi(\xi).b = \Phi(\xi.x_1) \quad \Phi(\xi).c = \Phi(\xi.y_0).
$$

**Remark 3.2.** The effects of doubling on continued fractions was first worked out by Hurwitz [9]. Raney introduced transducers for making calculations such as these in [10].

**Proof.** We will only prove the identities $\phi(\xi.y) = 2\phi(\xi)$ and $\Phi(\xi).c = \Phi(\xi.y_0)$; the remaining verifications are similar and left to the reader. We will first verify the identity $\phi(\xi.y) = 2\phi(\xi)$. Observe that, since $\phi$ and $y$ are continuous, it suffices to verify this equality for sequences which are eventually constant. The proof is now by induction on the minimum digit beyond which the sequence is constant. For the base case we have:

$$
\phi(\emptyset.y) = \phi(\emptyset) = 0 = 2 \cdot 0 \quad \phi(\bar{1}.y) = \phi(\bar{1}) = \infty = 2 \cdot \infty.
$$

In the inductive step, we have three cases:

$$
\phi(00\xi.y) = \phi(0\xi.y) = 1 + \frac{1}{1 + \phi(\xi.y)} = 1 + \frac{1}{2\phi(\xi)} = 2 + \frac{1}{\phi(\xi)} = 2\phi(00\xi)
$$

$$
\phi(01\xi.y) = \phi(10\xi.y^{-1}) = 1 + \frac{1}{1 + \phi(\xi.y^{-1})} = 1 + \frac{1}{2\phi(\xi)} = 1 + \frac{1}{1 + \phi(\xi)} = 2\phi(01\xi)
$$

$$
\phi(1\xi.y) = \phi(11\xi.y) = 2 + \phi(\xi,y) = 2 + 2\phi(\xi) = 2(1 + \phi(\xi)) = 2\phi(1\xi).
$$

Next we turn to the verification of $\Phi(\xi).c = \Phi(\xi.y_0)$. Observe that if $\xi$ does not extend 10, then $\Phi(\xi)$ is outside the interval $(0, 1)$ and we have $\Phi(\xi).c = \Phi(\xi) = \Phi(\xi.y_0)$. The remaining case follows from the identity we have already established, noting that $\frac{2t}{t + 1} = \frac{2}{t + 1/t}$:

$$
\Phi(10\xi).c = \left(1 + \frac{1}{\phi(\xi)}\right).c = \frac{2}{2 + \phi(\xi)} = 1 + \frac{1}{\phi(\xi)} = 1 + \frac{1}{\phi(\xi) - y} = \Phi(10\xi.y_0).
$$

□
From this point forward, we will identify \( a, b, \) and \( c \) with \( x, x_1, \) and \( y_{10} \), respectively and suppress all mention of \( \Phi \).

We now return to our discussion of the generators. It is straightforward to verify that the following relations are satisfied by these elements, where \( s \) and \( t \) are finite binary sequences:

1. \( x_s^2 = x_{s0}x_sx_{s1}; \)
2. if \( t.x_s \) is defined, then \( x_t.x_s = x_s.x_t.x_s; \)
3. if \( t.x_s \) is defined, then \( y.t.x_s = y.s.y.x_s; \)
4. if \( s \) and \( t \) are incompatible, then \( y.s.y.t = y.t.y.s; \)
5. \( y_s = x_s.y_{s0}y_{s10}y_{s11}. \)

We will refer to these relations collectively as \( R \). The first two groups of relations are known to give a presentation for \( F \): the function \( x_{1n} \) corresponds to the \( n \)th generator in the standard infinite presentation of \( F \). We will use \( F \) to denote the group generated by \( \{ x_s : s \in 2^{<\mathbb{N}} \} \).

Notice that any \( y_s \) is conjugate by an element of \( F \) to exactly one of \( y, y_0, y_1, \) or \( y_{10} \). Define \( X = \{ x_s : s \in 2^{<\mathbb{N}} \} \), \( Y = \{ y_s : s \in 2^{<\mathbb{N}} \} \), and \( Y_o \) to be the set of all \( y_s \) such that \( s \) is not a constant binary sequence. Observe that \( Y_o \) consists of those elements of \( Y \) which are conjugate to \( y_{10} \) by an element of \( F \). The group \( G_0 \) defined in the introduction is therefore generated by the (redundant) generating set \( S_0 = X \cup Y_o \).

Let \( R_0 \) be those relations in \( R \) which only refer to generators in \( S_0 \) and let \( G \) be the group generated by \( S = X \cup Y \). The rest of the paper will focus on proving the following theorem.

**Theorem 3.3.** The relations \( R \) give a presentation for \( G \) and the relations \( R_0 \) give a presentation for \( G_0 \). Moreover, \( G \) and \( G_0 \) admit finite presentations.

In the remainder of this section, we will prove that \( G \) and \( G_0 \) are finitely presented assuming that \( R \) and \( R_0 \) give presentations for these groups. The finite generating sets for these groups are \( \{ x, x_1, y_0, y_1, y_{10} \} \) and \( \{ a, b, c \} = \{ x, x_1, y_{10} \} \), respectively. Before proceeding, it will be necessary to define the other generators as words in terms of these generators; these definitions will be compatible with equalities which hold in \( G \). We begin by declaring

\[
y = x_{0}y_{10}y_{11}^{-1}, \quad x_{0} = x_{2}x_{1}^{-1}x^{-1}, \quad x_{10} = x_{1}^{2}x_{1}^{-1}x_{0}x_{1}^{-1}.
\]

Observe that \( 0.x^{-n} = 0^{n+1} \) and \( 1.x^{n} = 1^{n+1} \) and set

\[
x_{0n+1} = x^{n}x_{0}x^{-n}, \quad y_{0n+1} = x^{-n}y_{0}x^{-n}.
\]

If \( s \in 2^{<\mathbb{N}} \) is nonconstant, fix a word \( f_{s} \) in \( \{ x, x_1 \} \) such that \( 10.f_{s} = s \) and define

\[
x_{s} = f_{s}^{-1}x_{10}f_{s}, \quad y_{s} = f_{s}^{-1}y_{10}f_{s}.
\]

Next we note the following two standard properties of \( F \).

**Proposition 3.4.** If \( g \) is any element of \( F \) and \( s \) is a finite binary sequence such that \( s.g \) is defined, then \( x_{s}g = gx_{s}.g \). In particular if \( g, x_{s} \) and \( x_{s}.g \) are expressed as words in \( \{ x, x_1 \} \), then the above equality is derivable from the relations in (1) and (3) above.

**Proposition 3.5.** If \( u \leq_{\text{lex}} v \) are incompatible binary sequences, then there is a \( g \) in \( F \) and \( s \leq_{\text{lex}} t \) each of length at most 3 such that \( s.g = u \) and \( t.g = v \).
From these facts it follows that every relation in (1) is conjugate via an element of $F$ to a relation in (1) indexed by sequences of length at most 3. The relations in (5) are conjugate via elements of $F$ to a relation $y_s = x_s y_s^{-1} y_{s0} y_{s1}$ where $s \in \{0, 10, 1\}$. The relations in (6) can be expressed as $y_s g = g y_s$ where $s \in \{0, 10, 1\}$ and $g \in F$ such that $s g = s$. These can be derived from relations $y_s x_t = x_t y_s$ where $s, x_t$ and $s, t$ are binary sequences of length at most 3. In particular $G$ and $G_0$ are finitely presented. In the case of $G_0$, one can check that the following list of 9 relations actually suffice:

\[
\begin{align*}
x_1 x^{-2} x_1 x &= x^{-1} x_1 x x_1 x^{-1} \\
x_1 x^{-3} x_1 x^2 &= x^{-2} x_1 x^2 x_1 x^{-1} \\
y_{10} x_0 &= x_0 y_{10} \\
y_{10} x_{11} &= x_{11} y_{10} \\
y_{10} x_{111} &= x_{111} y_{10} \\
y_{10} y_{10} &= y_{10} y_{10} \\
y_{10} &= x_{10} y_{100} y_{1010} y_{1011}.
\end{align*}
\]

(Notice that all of the above relations except the last assert that a pair of elements of the group commute. In each case this is because they are supported on disjoint sets, where the support $g$ is the set of $x$ such that $x g \neq x$.) When expressed in terms of the original generators, these become:

\[
\begin{align*}
ba^{-2} ba &= a^{-1} baba^{-1} \\
ba^{-1} a^{-2} ba^2 &= a^{-2} ba^2 ba^{-1} \\
c a^{-1} b^{-1} a^{-1} &= a^2 b^{-1} a^{-1} c \\
c b^2 a^{-1} bab &= b^2 a^{-1} babc \\
ca^{-1} ba &= a^{-1} bac \\
ca^{-1} c &= a^{-2} ba^2 c \\
ca^{-2} a^{-1} c &= a^{-2} ca^{-2} c \\
c &= b^2 a^{-1} b^{-1} ab^{-1} c^{-1} ba^{-1} bab^{-1} ab^{-1} ca^{-1} ba^{-1}.
\end{align*}
\]

4. Tree diagrams

Before proceeding further, we will pause to describe how the elements of $\langle a, b, c \rangle$ can be described in terms of tree diagrams, similar to those associated to Thompson’s group $F$. This section is not essential for understanding the proof of Theorem 1.1 in the subsequent section, although the reader may find the material here is useful in visualizing what is happening in the main proofs.

Let $\mathcal{T}$ denote the collection of all finite sets $S$ of reduced words in the alphabet $\{0, 1, y, y^{-1}\}$ which satisfy the following properties:

- $S$ is nonempty;
- the result of deleting all occurrences of $y$ and $y^{-1}$ in members of $S$ defines a bijective map between $S$ and an element of $\mathcal{T}$;
- if $uy^n$ is a prefix of some element of $S$, then any element of $S$ which has $u$ as a prefix, also has $uy^n$ as a prefix;

Elements of $\overline{T}$ be be visualized as follows. Let $S$ be in $\mathcal{T}$ and $T$ is the result of removing the occurrences of $y$ and $y^{-1}$ from elements of $S$. We can think of $T$ as defining a rooted ordered binary tree, whose vertexes correspond to the prefixes of elements of $T$. The elements of $S$ can be specified by an assignment of an integer to each vertex of $T$. For instance if $S = \{0, 1yy0y^{-1}, 1yy1\}$, then the associated labeled tree is:

```
-1
  2
```
A labeled tree diagram is a pair $S \to T$ of elements of $\tilde{T}$ such that $S$ and $T$ have the same number of vertices. The key point is now to define the appropriate notion of equivalence of tree diagrams. First we define a notion of equivalence on $\tilde{T}$. Two (possibly infinite) words in the alphabet $\{0, 1, y, y^{-1}\}$ are equivalent if one can be converted into the other by the following substitutions:

- $y00 \iff 0y$
- $y01 \iff 10y^{-1}$
- $y1 \iff 11y$
- $y^{-1}1 \iff 11y^{-1}$
- $y^{-1}10 \iff 01y$
- $y^{-1}11 \iff 1y$.

Two elements of $\tilde{T}$ are equivalent if the sets of equivalence classes of their elements coincide. In terms of labeled tree diagrams, this means that $S$ and $T$ are equivalent if $T$ can be obtained from $S$ by a sequence of substitutions of the following form:

\[
\begin{array}{c}
\begin{array}{c}
\text{m} \\
\text{i} \\
\end{array} & \quad \\
\begin{array}{c}
\text{k} \\
\end{array} \\
\end{array}
\quad \iff 
\begin{array}{c}
\begin{array}{c}
\text{m-1} \\
\text{k+1} \\
\end{array} \\
\begin{array}{c}
\text{j-1} \\
\end{array} \\
\text{i+1}
\end{array}
\]

In many simple computations, labels are either 0, 1 or $-1$. In this case it is convenient to use $\bullet$ and $\circ$ to indicate the labels 1 and $-1$ respectively. The substitution rule above then becomes a pair of substitutions:

\[
\begin{array}{c}
\begin{array}{c}
\text{m} \\
\text{i} \\
\end{array} & \quad \\
\begin{array}{c}
\text{k} \\
\end{array} \\
\end{array}
\quad \iff 
\begin{array}{c}
\begin{array}{c}
\text{m-1} \\
\text{k+1} \\
\end{array} \\
\begin{array}{c}
\text{j-1} \\
\end{array} \\
\text{i+1}
\end{array}
\]

Notice that if $S$ and $T$ are equivalent elements of $\tilde{T}$, then $S$ and $T$ have the same number of leaves.

Equivalence of labeled tree diagrams is generated by the equivalence of trees, together with the following manipulations on tree diagrams:

- If $S \to T$ is a labeled tree diagram, then we can insert a caret below the $i^{th}$ leaf of $S$ and below the $i^{th}$ leaf of $T$ to produce an equivalent diagram $S' \to T'$. The labels of the top vertexes of the new carets are the same as the original vertexes; the leaves of the new carets are labeled 0.
- If $S \to T$ is a labeled tree diagram, then we may add 1 to the label of the $i^{th}$ leaves of $S$ and of $T$ to produce an equivalent diagram $S' \to T'$.

If $S \to T$ is a labeled tree diagram and $S$ has no labels, then it describes a continuous function $g : 2^\mathbb{N} \to 2^\mathbb{N}$ as follows. If $\xi$ is an infinite sequence in the alphabet $\{0, 1, y, y^{-1}\}$ with only finitely many occurrences of $y$ or $y^{-1}$, then define $\lim \xi$ to be the unique infinite binary sequence $\eta$ such that every prefix of $\eta$ occurs as a prefix of a sequence equivalent to $\xi$. If $s_i$ and $t_i$ are the $i^{th}$-least elements of $S$ and $T$ respectively in the lexicographic order, then define $g(s_i\xi) = \lim t_i\xi$. It is easy to check that the generators can be described as follows:

\[
a = ( \begin{array}{c}
\cdot \\
\cdot \\
\end{array} \to \begin{array}{c}
\cdot \\
\cdot \\
\end{array} ) \quad b = ( \begin{array}{c}
\cdot \\
\cdot \\
\end{array} \to \begin{array}{c}
\cdot \\
\cdot \\
\end{array} ) \quad c = ( \begin{array}{c}
\cdot \\
\cdot \\
\end{array} \to \begin{array}{c}
\cdot \\
\cdot \\
\end{array} )
\]

In fact we can modify this definition slightly in order to associate a function to any labeled tree diagram: define $g(\lim s_i\xi) = \lim t_i\xi$. We leave it to the reader to verify that this is a well defined map. The equivalence of tree diagrams is set up so as to capture exactly when the corresponding functions coincide. We will eventually see that the collection of all functions arising in this way is a group which then coincides with the group $G$ of the previous section.
functions associated to these diagrams is the same as that described by \( S \to U \). In particular \( T \to S \) is the inverse of \( S \to T \).

We will conclude this section with a illustrative computation. Notice that \( t \mapsto 2t \) correspond to the diagram \( \wedge \to \sigma \bullet \). Conjugating \( t \mapsto t + 1 \) by \( t \mapsto 2t \) yields \( t \mapsto t + 2 \), the square of the first map. In terms of labeled tree diagrams, this computation can be carried out as follows:

\[
\left( \wedge \to \sigma \bullet \right)^{-1} \cdot \left( \wedge \to \sigma \bullet \right) \cdot \left( \wedge \to \sigma \bullet \right) =
\]

\[
\left( \wedge \to \bullet \right) \cdot \left( \wedge \to \bullet \right) \cdot \left( \wedge \to \bullet \right) =
\]

\[
\bullet \to \bullet = (\left( \wedge \to \bullet \right)^2)
\]

5. Sufficiency of the relations

In this section, we will prove that the relations in \( R \) and \( R_0 \) are sufficient to give presentations for \( G \) and \( G_0 \). We will use without proof that the relations in \( R \) which only refer to the generators in \( X \) give a presentation for \( F \) (see [3] [5]). The strategy of the proof is as follows. First, we will argue that any \( S \)-word can be put into a standard form by applying the relations. Standard forms are not unique, but are organized in a way which better facilitates further symbolic manipulations. We will then define the notion of a sufficiently expanded standard form, argue that every standard form can be sufficiently expanded by applying the relations in \( R \), and that any sufficiently expanded standard form which represents an element of \( F \) is an \( X \)-word.

We will begin by defining some terminology. In what follows, we will say that an \( S \)-word \( \Omega_1 \) is derived from an \( S \)-word \( \Omega_0 \) if it is the result of applying substitutions of the following forms:

\[
y^i s^\pm 1 \Rightarrow x^\pm 1 t x^\pm 1 \quad y_s \Rightarrow x_s y_o y_{s10} y_{s11}
\]

\[
y_u y_v \Leftrightarrow y_v y_u \quad x^{i+j} \Leftrightarrow x^i x^j \quad y^{i+j} \Leftrightarrow y^i y^j
\]

\[
delete an occurrence of y^j y^{-i}
\]

where \( s, t, u, v \in 2 \leq \mathbb{N} \) are such that \( t x_s \) is defined and \( u \) and \( v \) are incompatible, and \( i, j \) are nonzero integers of the same sign. We will write this symbolically as \( \Omega_0 \Rightarrow \Omega_1 \). Notice that each of these substitutions corresponds either to a relation in \( R \) or to a group-theoretic identity. Also observe that only \( S_0 \)-words can be derived from \( S_0 \)-words.

**Definition 5.1.** An \( S \)-word \( \Omega \) is in standard form if it is the concatenation of a \( X \)-word followed by a \( Y \)-word and whenever \( \Omega(i) = y^m \), \( \Omega(j) = y^n \), and \( s \subseteq t \), then \( j \leq i \). We will write standard form to mean an \( S \)-word in standard form. The depth of a standard form \( \Omega \) is the least \( l \) such that there is binary sequence \( s \) of length \( l \) such that \( y_s \) occurs in \( \Omega \) (if \( \Omega \) is an \( X \)-word, then we say that \( \Omega \) has infinite depth).

Notice in particular that a given \( y_s \) can occur at most once in a standard form (although possibly with an exponent other than \( \pm 1 \)). Observe that any group element which is expressible by a word in standard form allows us to describe the group element via a labeled tree diagram in the sense of the previous section.
Lemma 5.2. For every \( s \in 2^{\mathbb{N}} \) and every \( l \in \mathbb{N} \), there is a standard form \( \Omega \) which can be derived from \( y_s^{l+1} \) such that:

1. if \( x_u \) occurs in \( \Omega \), then \( u \) extends \( s \);
2. if \( y_u \) occurs in \( \Omega \), then \( u \) extends \( s \), has length at least \( l \), and the exponent of \( y_u \) is \( \pm 1 \);
3. if \( y_u \) and \( y_v \) occur in \( \Omega \) and \( u \neq v \), then \( u \) and \( v \) are incompatible.

Proof. The proof is by induction on \( l - |s| \). If \( l - |s| = 0 \), there is nothing to do since \( y_s \) already satisfies the conclusion of the lemma. If \( l - |s| > 0 \), then \( y_s \Rightarrow x_s y_0 y_0 y_1 y_1 \) and we can apply the induction hypothesis to obtain \( \Omega_{s_0}, \Omega_{s_1}, \Omega_{s_1} \) which satisfy the conclusion of the lemma for \( y_{s_0}, y_{s_1}, y_{s_1} \) respectively for the same value of \( l \). By conclusion 3 of the lemma, we can apply substitutions of the form \( y_v x_u \Rightarrow x_u y_v \) for incompatible \( u \) and \( v \) move the occurrence of \( x_u \) in

\[ x_s \Omega_{s_0} \Omega_{s_1} \Omega_{s_1} \]

to the left, placing the word a standard form which satisfies the conclusions of the lemma. The case of \( y_s^{-1} \) is handled similarly using the substitution \( y_s^{-1} \Rightarrow x_s^{-1} y_{s_0} y_{s_1} y_{s_1}^{-1} \).

Lemma 5.3. If \( \Xi \) is an \( X \)-word, then there is an \( l_0 \) such that if \( \Omega \) is a standard form of depth \( l \geq l_0 \), then \( \Omega \Xi \Rightarrow \Omega' \) for some standard form \( \Omega' \) of depth at least \( l - k \) where \( k \) is the word length of \( \Xi \).

Proof. If \( \Xi = x_s^{l+1} \), then observe that if \( t \) is a finite binary sequence of length at least \( l + 2 \), then \( t x_s^{l+1} \) is defined and its length differs from the length of \( \Omega \) by at most 1. Thus by repeated applications of the substitution \( y_t' \Rightarrow x_s^{l+1} y_{t,x_s^{l+1}}' \), the final occurrence of \( x_s \) in \( \Omega^{l+1} x_s \) can be moved to the left of all occurrences of a \( y_t \). This results in a standard form in which the depth is changed by at most 1. The general case now follows by induction.

Lemma 5.4. If \( \Omega \) is any \( S \)-word and \( l \in \mathbb{N} \), then \( \Omega \Rightarrow \Omega' \) for some standard form \( \Omega' \) of depth at least \( l \).

Proof. The proof is by double induction: first on the word length \( n \) of \( \Omega \) and then on \( l \). The case \( n = 1 \) is handled by Lemma 5.2. By making a substitution of the form \( a^{k+1} \Rightarrow a^{k+1} \) if necessary, we may assume that \( \Omega = \Omega_0 \Omega_1 \) where \( \Omega_1 \) is an \( S \)-word of positive length. By our induction hypothesis \( \Omega_1 \Rightarrow \Xi \), where \( \Xi \) is an \( X \)-word respectively and \( \Xi \) has depth \( l \). Let \( k \) be the word length of \( \Xi \) and let \( m \geq l \) be such that if \( y_u \) occurs in \( \Xi \), \( u \) has length less than \( m \). By our induction hypothesis, there is a standard form \( \Omega'_0 \) of depth at least \( m + k \) such that \( \Omega_0 \Rightarrow \Omega'_0 \). By Lemma 5.2 we have that \( \Omega'_0 \Xi \Rightarrow \Omega''_0 \Xi \) for some standard form \( \Omega''_0 \) of depth at least \( m \), we have:

\[ \Omega \Rightarrow \Omega_0 \Omega_1 \Rightarrow \Omega_0 \Xi \Rightarrow \Omega'_0 \Xi \Rightarrow \Omega''_0 \Xi \]

Finally, notice that since the depth of \( \Omega''_0 \Xi \) is at least \( m \), \( \Omega' = \Omega''_0 \Xi \) is a standard form of depth at least \( l \), as desired.

If \( \Omega \) is standard form and \( y_s \) occurs in \( \Omega \), we say that \( s \) is exposed in \( \Omega \) if there is a finite binary sequence \( u \) extending \( s \) such that if \( t \) is a binary sequence compatible with \( u \) and \( y_t \) occurs in \( \Omega \), then \( t \) is an initial part of \( s \).
Definition 5.5. A standard form \( \Omega \) is **sufficiently expanded** if whenever \( y_s \) occurs in \( \Omega \) and \( s \) is not exposed in \( \Omega \), then:

- \( y_{s0} \) occurs in \( \Omega \) if \( y_s \) occurs positively in \( \Omega \);
- \( y_{s1} \) occurs in \( \Omega \) if \( y_s \) occurs negatively in \( \Omega \).

The motivation for this definition is as follows. Suppose that \( \Omega \) is a standard form which is not sufficiently expanded and that this is witnessed by \( \Omega(i) = y_s^n \) for \( n > 0 \). If we substitute

\[
x_s y_{s0} y_{s10} y_{s11} y_s^{n-1}
\]

for \( y_s^n \) in \( \Omega \), then whenever \( y_t \) occurs before \( y_s \) in \( \Omega \), \( t.x_s \) is defined. A similar conclusion holds — with \( x_s^{-1} \) replacing \( x_s \) — if \( n < 0 \) and the substitution

\[
x_s^{-1} y_{s0} y_{s10} y_{s11} y_s^{n+1}
\]

is applied. This plays an important role in the proof of the next lemma.

**Lemma 5.6.** If \( \Omega \) is a standard form, then there is a sufficiently expanded standard form which can be derived from \( \Omega \).

**Proof.** We will prove the lemma by defining a well-founded partial ordering \( \prec \) on the set of standard forms and a notion of expansion on standard forms which are not sufficiently expanded in such a way that produces a smaller standard form in this ordering. Here a partial order is well-founded if it has no infinite strictly decreasing sequences. We will define the ordering first.

If \( \Omega \) is an \( S \)-word, let \( T(\Omega) \) denote the minimal prefix set which has the property that if \( y_t \) occurs in \( \Omega \), then \( t \) has an extension in \( T \). If \( \Omega_0 \) and \( \Omega_1 \) are standard forms, define \( \Omega_0 \prec \Omega_1 \) if \( |T(\Omega_0)| < |T(\Omega_1)| \) or \( |T(\Omega_0)| = |T(\Omega_1)| \) and \( |k_0| < |k_1| \) where \( k_i \) is the exponent in \( \Omega_i \) of the \( \leq_{\text{lex}} \)-maximal \( s \) such that \( y_s \) occurs in at least one of \( \Omega_0 \) or \( \Omega_1 \) and for which \( k_0 \neq k_1 \) (if \( y_s \) occurs in only one of the \( \Omega_i \), then the other exponent is 0). Notice that for a fixed \( m \) there are only finitely many prefix sets of cardinality \( m \). In particular the collection \( F \) of all finite binary sequences which have an extension in a prefix set of cardinality \( m \) is finite. Since the lexicographic ordering on \( \mathbb{N}^F \) is a well-order, \( \prec \) is well-founded.

Now suppose that \( \Omega \) is a standard form which is not sufficiently expanded as witnessed by \( \Omega(i) = y_s^n \). For simplicity, suppose that \( n > 0 \) and apply the substitution

\[
y_s^n \Rightarrow x_s y_{s0} y_{s10} y_{s11} y_s^{n-1}
\]

(if \( n = 1 \), the \( y_s^{n-1} \) term is omitted) followed by substitutions of the form \( y_t^m x_s \Rightarrow x_s y_t^m x_s \), to move \( x_t \) to the left, forming a new word \( \Omega' \) which is the concatenation of a \( X \)-word followed by a \( Y \)-word. At this point, the only thing preventing \( \Omega' \) from being a standard form is the newly introduced occurrences of \( y_{s0}, y_{s10}, \) and \( y_{s11} \). Observe that \( y_{s1} \) can not occur in \( \Omega' \); for this to happen, \( s1 \) would have to equal \( t.x_s \) for some \( t \) such that \( t.x_s \) is defined, and such a \( t \) does not exist. Furthermore, if \( y_t \) occurs in \( \Omega' \) and \( t \) properly extends one of the sequences \( s0, s10, \) or \( s11 \), then \( y_t \) must occur before any occurrence of \( y_{s0}, y_{s10}, \) or \( y_{s11} \) in \( \Omega' \). Similarly, if \( t \) is a proper initial part of \( s0, s10, \) and \( s11 \) and \( y_t \) occurs in \( \Omega' \), then \( t \) is actually an initial part of \( s \) and thus the occurrence is after the point of the substitution. We may therefore apply substitution of the form \( y_t y_v \Rightarrow y_v y_u \) for incompatible \( u \) and \( v \) in order to move any two distinct occurrences of \( y_{s0}, y_{s10}, \) or \( y_{s11} \) to the same position in \( \Omega' \), resulting in a word \( \Omega'' \) which is now in standard form.
Lemma 5.9. Suppose that \( T \) is a string of elements of \( \Lambda \). It follows that \( t.x_t \) is defined for every element \( t \) of \( T(\Omega) \) and that \( T(\Omega') = \{ t.x_t : t \in T(\Omega) \} \). Hence \( T(\Omega) \) and \( T(\Omega') \) have the same cardinality. Notice that if \( y_s \) occurs in \( \Omega'' \), it occurs in \( \Omega' \) and hence \( T(\Omega') \) dominates \( T(\Omega'') \). It follows that \( T(\Omega'') \) has cardinality at most that of \( T(\Omega) \). If \( T(\Omega'') \) has the same cardinality as \( T(\Omega) \), then \( s \) is the \( \leq_{\text{lex}} \)-maximal sequence such that the exponent of \( y_s \) in \( \Omega \) and \( \Omega'' \) differs and in this case, it decreases by one in absolute value. Thus we have shown \( \Omega'' \triangleleft \Omega \).

Let \( B \) denote the set \{0, 1, y, y^{-1}\} and let \( B^{<N} \) denote the collection of all finite strings of elements of \( B \). If \( \Lambda \) is in \( B^{<N} \) and \( \Lambda(i) \) is either \( y \) or \( y^{-1} \), we will say that \( \Lambda(i) \) is an occurrence of \( y^\pm \). We will use \( B \)-words to analyze the evaluation of standard forms at binary sequences. The following symbolic manipulations correspond to the recursive definition of the function \( y : 2^N \rightarrow 2^N \).

**Definition 5.7.** Suppose that \( \Lambda \) is in \( B^{<N} \). An application of one of the substitutions

\[
\begin{align*}
00 & \Rightarrow y0y \\
y00 & \Rightarrow 0y \\
y01 & \Rightarrow 10y^{-1}
\end{align*}
\]

at an occurrence of \( y^\pm \) is said to advance that symbol. If several advances of occurrences of \( y^\pm \) are applied to \( \Lambda \), resulting in \( \Lambda' \), then we say that \( \Lambda \) can be advanced to \( \Lambda' \), denoted \( \Lambda \Rightarrow \Lambda' \).

**Definition 5.8.** Suppose that \( \Lambda \) is in \( B^{<N} \). An occurrence of \( y^\pm \) is a potential cancellation in \( \Lambda \) if repeatedly advancing it results in an occurrence of the substring \( yy^{-1} \) or \( y^{-1}y \) in the modified word.

**Lemma 5.9.** Suppose that \( \Lambda \) is in \( B^{<N} \) and contains no potential cancellations. Then advancing any occurrence of a \( y^\pm \) results in a word with no potential cancellations.

**Proof.** Suppose that \( \Lambda \) is given and that the \( i \)th occurrence of \( y^\pm \) is advanced to \( \Lambda' \). The only possibility for introducing a potential cancellation is if \( i > 1 \) and a potential cancellation occurs at the \( i - 1 \)st occurrence of a \( y^\pm \) in \( \Lambda' \). Return to \( \Lambda \) and advance the \( i - 1 \)st occurrence of \( y^\pm \) as much as possible to produce \( \Lambda'' \).

Suppose for a moment that after advancing, this occurrence is a \( y \); notice that the next symbol is either \( 0 \) or \( y \). We now have the following cases:

\[
\begin{align*}
yy00 & \Rightarrow yy0 \\
y000 & \Rightarrow 0y0y \Rightarrow y000y \Rightarrow 0y0y \\
y001 & \Rightarrow 0y0y \Rightarrow 0y0y \Rightarrow 0y0y \\
y0010 & \Rightarrow 10y^{-1}0y^{-1} \Rightarrow 1000y^{-1}y^{-1} \\
y0011 & \Rightarrow 10y^{-1}1y \\
y0y00 & \Rightarrow 0y00y \Rightarrow 0y00y \\
y0y01 & \Rightarrow 0y00y \Rightarrow 0y00y \\
y0y10 & \Rightarrow 0y00y \Rightarrow 0y00y \\
y0y11 & \Rightarrow 0y00y \\
y0y^{-1}0 & \Rightarrow 0y00y \\
y0y^{-1}10 & \Rightarrow 0y00y \\
y0y^{-1}11 & \Rightarrow 0y00y \\
y0y^{-1}00 & \Rightarrow 0y00y \\
y0y^{-1}01 & \Rightarrow 0y00y \\
y0y^{-1}10 & \Rightarrow 0y00y \\
y0y^{-1}11 & \Rightarrow 0y00y
\end{align*}
\]
The above lines list the possible contexts for two occurrences of $y^\pm$ in the $i - 1$st and $i$th in $\Lambda''$ where the first occurrence is positive. In the above cases, the $i$th occurrence of $y^\pm$ is advanced in $\Lambda''$ and then the $i - 1$st occurrence is advanced as much as possible, demonstrating that a cancellation does not occur. The case in which the $i - 1$st occurrence of $y^\pm$ in $\Lambda''$ is $y^{-1}$ is handled by symmetry — the rules for advancement and potential cancellation are invariant under the following involution:

\[ y \leftrightarrow y^{-1} \quad 0 \leftrightarrow 1. \]

**Lemma 5.10.** Suppose that $\Lambda$ is in $B^{<\aleph}$ and contains no potential cancellations. There is a finite binary sequence $u$ such that $\Lambda^\sim u$ can be advanced to $s^\sim y^n$ for some binary sequence $s$, where $n$ is the number of occurrences of $y^\pm$ in $\Lambda$.

**Proof.** The proof is by induction on the number of occurrences of $y^\pm$ in $\Lambda$. If there is only one occurrence, advance the occurrence as many times as possible, resulting in $s^\sim y$, $s^\sim y^{-1}$, $s^\sim y^0$, or $s^\sim y^{-1}1$ for some finite binary sequence $s$. In the first case we are finished; in the remaining cases, the choices $u = 10$, $u = 0$, and $u = 0$ work. Now suppose that $\Lambda$ contains $n + 1$ occurrences of $y^\pm$. Induction and Lemma 5.9 reduce the general case to the two special cases $y0y^n$ and $y^{-1}1y^n$. In these cases, use $u = 0^n0^n$, observing that $y^n02^n$ can be advanced to $0y^n$. □

**Lemma 5.11.** If $\Omega$ is a sufficiently expanded standard form then either $\Omega$ is an $X$-word or else $\Omega$ does not have the same evaluation as an $X$-word.

**Proof.** Notice that it is sufficient to prove the lemma when $\Omega$ is a sufficiently expanded standard form which is a $Y$-word of positive length. Suppose that such an $\Omega$ is given and let $g : 2^\aleph \rightarrow 2^\aleph$ be the evaluation of $\Omega$ in $G$. It will be sufficient to construct finite binary sequences $u$ and $v$ such that the value of $g$ at $u^\sim \xi$ is $v^\sim y^n(\xi)$ for some $n > 0$. This is because if $\xi = 0^2102^n1\ldots$, then the value of $g$ at $u^\sim \xi$ is $v^\sim 0^12^n01^2\ldots$, which is not tail equivalent to $u^\sim \xi$.

The finite binary sequence $u$ will be constructed by a recursive procedure. Let $u \upharpoonright i_0$ be the finite binary sequence such that the last entry of $\Omega$ is a power of $y_{i_0}$. Suppose that $u \upharpoonright i$ has been defined and that $y_{u|i}$ occurs in $\Omega$. If $u \upharpoonright i$ is exposed in $\Omega$, then let $u \upharpoonright i + 1$ be any finite binary sequence extending $u \upharpoonright i$ which witnesses this. Otherwise, define $u(i) = 0$ if $y_{u|i}$ occurs positively in $\Omega$ and $u(i) = 1$ if $y_{u|i}$ occurs negatively in $\Omega$. Define $\Lambda$ to be the result of simultaneously inserting $y^n$ after $s \upharpoonright i$ whenever $y^n$ occurs in $\Omega$. Notice that by the choice of our sequence $u$, $\Lambda$ does not contain potential cancellations: any occurrence of $y$ except for the final occurrence of $y^\pm$, is followed by $0y^\pm$ and any occurrence of $y^{-1}$ except for the final occurrence of $y^\pm$ is followed by $1y^\pm$. It follows from Lemma 5.10 that there is a sequence $s$ such that $\Lambda^\sim s$ can be advanced to $vy^n$ for some binary sequence $v$, where $n$ is the number of occurrences of $y^\pm$ in $\Lambda$ (this number coincides with the number of steps of the recursive procedure above, which is at least 1). Set $u$ to be the concatenation of $u \upharpoonright i + 1$ followed by $s$.

Recall now that we have assumed that $\Omega$ is a $Y$-word; $g = y_{t_{ik}}^{n_k} \ldots y_{t_1}^{n_1}$.

\[ \xi \cdot g = \xi.(y_{t_i}^{n_i} \ldots y_{t_1}^{n_1}) = (\cdots (\xi, y_{t_2}^{n_2}) \cdots) \cdot y_{t_k}^{n_k} \]

Let $\xi_i$ be the result of applying the $y_{t_i}^{n_i} \ldots y_{t_{i-1}}^{n_{i-1}}$ to $\xi$. Observe that if $t_{i+1}$ is an initial part of $\xi_i$, then it is still an initial part of $\xi_i$. This follows from the fact
that if $j \leq i$, then $t_j$ either extends $t_{i+1}$ or else is incompatible with $t_{i+1}$. In particular, if $t_{i+1}$ is not an initial part of $\xi$, then $\xi_{i+1} = \xi_i$. If $t_{i+1}$ is an initial part of $\xi_i$, then $\xi_{i+1} = t_{i+1} \mathbin{\cdot} \eta_i$ (where $\xi_i = t_i \mathbin{\cdot} \eta_i$). It follows from $A \Rightarrow v y^n$ that $\xi. y = v^n (\eta. y^n)$, where $\eta$ is such that $\xi = u \mathbin{\cdot} \eta$.

To see that this finishes the proof of the main theorem, suppose that $\Omega$ is an $S$-word which evaluates to the identity function in $G$. By Lemma 5.7, $\Omega \Rightarrow \Omega'$ for some word $\Omega'$ in standard form. By Lemma 5.6, $\Omega' \Rightarrow \Omega''$ for some word $\Omega''$ which is in standard form and which is sufficiently expanded. In particular, $\Omega''$ is equivalent to $\Omega$ by the relations in $R$; if $\Omega$ was an $S_0$-word, then $\Omega''$ is an $S_0$ word and the derivation $\Omega \Rightarrow \Omega' \Rightarrow \Omega''$ utilizes only relations in $R_0$. By Lemma 5.11, $\Omega''$ is an $X$-word. Since $R$ includes a presentation for $F$, $\Omega''$ can be reduced to the identity using the relations in the relations in $R_0$.

REFERENCES

[1] S.I. Adyan. The Burnside Problem and Identities in Groups. volume 95 of Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer, Berlin, 1979.

[2] S.I. Adyan. Random walks on free periodic groups. Izv. Akad. Nauk SSSR Ser. Mat. 46(6):1139–1149, 1982.

[3] J. Belk. Thompson’s group $F$. Ph.D. Thesis, Cornell University, 2004. [arXiv:0708.3609]

[4] M.G. Brin, C.C. Squier. Groups of piecewise linear homeomorphisms of the real line. Invent. Math., 79(3):485–498, 1985.

[5] J.W. Cannon, W.J. Floyd, W.R. Parry. Introductory notes on Richard Thompson’s groups. Enseign. Math. (2), 42(3-4):215–256, 1996.

[6] Y. Carrière, É. Ghys. Relations d’équivalence moyennables sur les groupes de Lie. C. R. Acad. Sci. Paris Sér. I Math, 300(19):677–680, 1985.

[7] A. Connes, J. Feldman, B. Weiss. An amenable equivalence relation is generated by a single transformation. Ergodic Theory Dynamical Systems, 1(4):431–450, 1982.

[8] M.M. Day. Means for the bounded functions and ergodicity of the bounded representations of semi-groups. Trans. Amer. Math. Soc., 69:276–291, 1950.

[9] A. Hurwitz. Über die Kettenbruch-Entwicklung der Zahl $e$. Phys.-6kon. Ges., Königsberg, 1891. Mathematische Werke, 2:129—133, Basel: Birkhäuser, 1933.

[10] S.V. Ivanov. Embedding free Burnside groups in finitely presented groups. Geommetriae dedicata, 111:87–105, 2005.

[11] A.S. Kechris, B.D. Miller, Topics in Orbit Equivalence. volume 1852 of Lecture Notes in Mathematics Springer, Berlin, 2004.

[12] N. Monod. Groups of piecewise projective homeomorphisms. Proc. Natl. Acad. Sci., 110(12):4524–4527, 2013

[13] A.Y. Ol’shanskii. On the question of the existence of an invariant mean on a group. Uspekhi Mat. Nauk, 35(4):199–200, 1980.

[14] A.Y. Ol’shanskii, M.V. Sapir. Non-amenable finitely presented torsion-by-cyclic groups. Publ. Math. Inst. Hautes Études Sci., 96:43–169, 2003.

[15] A.L.T. Paterson. Amenability, volume 29 of Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 1988.

[16] G.N. Raney. On continued fractions and finite automata. Math. Ann., 206:265–283, 1973.

[17] J. von Neumann. Zur allgemeinen Theorie des Maßes. Fund. Math., 13:73–116, 1929.

[18] J. Zimmer. Amenable ergodic group actions and an Application to Poisson boundaries of random walks. Journal of Functional Analysis, 27:350–372, 1978.

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NY 14853–4201, USA

E-mail address: yl763@cornell.edu