NON-FINITELY GENERATED MAXIMAL SUBGROUPS OF CONTEXT-FREE MONOIDS

CARL-FREDRIK NYBERG-BRODDA

Abstract. A finitely generated group or monoid is said to be context-free if it has context-free word problem. In this note, we give an example of a context-free monoid, none of whose maximal subgroups are finitely generated. This answers a question of Brough, Cain & Pfeiffer on whether the group of units of a context-free monoid is always finitely generated, and highlights some of the contrasts between context-free monoids and context-free groups. Finally, we ask whether the group of units of a context-free monoid is always coherent.

1. Introduction

The study of language-theoretic properties of algebraic structures can be seen as a natural extension of the work by Thue \cite{thue1901} in the early 20th century on the word problem (as a decision problem) for monoids. Thue, working with what today are called rewriting systems (or semi-Thue systems) achieved a remarkable number of results that are today easily recognisable; for example, he formulated the word problem for finitely presented monoids, and solved this problem for certain one-relation monoids. This can be seen as a first step in exploring the connection between rewriting of words – and, more generally, formal language theory – and algebraic structures. Anisimov \cite{anisimov1971} significantly strengthened this connection in 1971. For a finitely generated group $G$, he studied the set of words over a given finite generating set equal to the identity element in $G$, calling it the word problem of $G$. He showed that this language is regular if and only if $G$ is finite. This was extended in 1983 by the remarkable theorem of Muller & Schupp \cite{muller1983}, who proved (when supplemented by a result of Dunwoody’s \cite{dunwoody1983}) that the word problem for $G$ is a context-free language if and only if $G$ is virtually free, i.e. $G$ has a free subgroup of finite index. The study of the word problem of groups for other classes of languages has expanded in many directions since, see. e.g. \cite{schutzenberger1963, schutzenberger1965, grigorchuk1975}.

Inspired by this success, Duncan & Gilman \cite{duncan1994} \S5 introduced a generalisation of the word problem to finitely generated monoids. This generalisation is not a direct translation; indeed, the language of words equal to the identity element in a monoid is, generally speaking, completely insufficient to describe the structure of the monoid (for notable exceptions to this, see \cite{grigorchuk1975}). However, their generalisation, which we shall describe in \S2 has the same language-theoretic properties as the original definition in the case that the monoid in question is a group; furthermore, its properties does not generally depend on the finite generating set chosen for the monoid. This generalisation also has the pleasing property that a monoid has

\begin{flushright}
\textbf{Date:} November 2, 2021.
\end{flushright}

2020 Mathematics Subject Classification. 20F10 (primary) 08A50, 20F05, 68R15, 68Q42.

The author gratefully acknowledges funding from the Dame Kathleen Ollerenshaw Trust, which is supporting his research at the University of Manchester.
regular word problem if and only if it is finite; that is, the analogue for monoids of Anisimov’s theorem holds.

However, the structure of monoids with context-free word problem seems at present hopelessly difficult to classify in full generality. Any theorem along the lines of the Muller-Schupp theorem for monoids rather than groups would have to be exceptionally deep indeed. One initial hurdle is that there is not even a natural notion of what a “finite index submonoid” of a monoid should be; or rather, perhaps, there are too many, none of which is quite satisfactory for these purposes (see e.g. [28, 41, 20, 21]). Even answering the question of what a “virtually free” monoid should be is hence already a non-trivial task. On the other hand, some less ambitious facts about context-free monoids are known; for example, one can easily show that free monoids have context-free word problem, and recently success has been found regarding closure properties of the class of context-free monoids [8].

The study of semigroup and monoid properties inherited from and by their maximal subgroups has a seen a good deal of attention in the literature. For some examples, see e.g. [17, 14, 43, 18, 19]. This study has also seen success regarding formal language theory. For example, the author has shown that the language-theoretic properties of the word problem of a finitely presented special monoid (see §2) are completely determined by its group of units [40]. Here the group of units of a monoid $M$ is the maximal subgroup containing the identity element 1 of $M$. The author has also proved analogous results for subspecial and compressible monoids [39]. Thus it is also natural to ask to what extent the word problem (in the above sense) of a monoid relates to the word problems of its maximal subgroups.

In line with this, Brough, Cain & Pfeiffer posed the question of whether the group of units of a context-free monoid is necessarily finitely generated. In this note, we answer this question negatively. In fact, we show something stronger.

**Theorem A.** There exists a context-free monoid none of whose maximal subgroups are finitely generated.

As context-free groups and monoids are, as part of their definition, finitely generated, it follows further that a maximal subgroup of a context-free monoid need not be context-free itself. This highlights the fact that a general reduction of the properties of context-free monoids to those of context-free groups seems difficult.

We remark that the monoid with the above properties constructed in this note is a non-finitely presented special monoid. Strong structural results are known for finitely presented special monoids via Makanin’s work on such monoids [32, 33], see also Zhang [51]. However, their results are stated only in the finitely presented case, and though much of their analysis carries over to the non-finitely presented case, we have, in the interest of remaining self-contained, opted to give direct proofs of all necessary statements. These proofs should also serve to give a taste of the flavour of many of the arguments arising in the theory of special monoids and combinatorial monoid theory. We remark also that some results for finitely presented special monoids fail in the non-finitely presented case; for example, one can show that a finitely presented special monoid $M$ has context-free word problem if and only if its group of units does [40]. This equivalence no longer holds if the assumption

---

1Indeed, A. K. Sushkevich, one of the fathers of semigroup theory, considered a problem about semigroups to be solved “if it could be reduced to a problem about groups” [46, §38].

2This appears as Question 10.4 in a pre-print version of [8], available at arXiv:1903.10493, but not in the final published version.
of finite presentability is dropped; in fact, there exists a finitely generated special monoid whose word problem is not context-free, and yet having trivial (and hence context-free) group of units [41].

2. Background

We assume the reader is familiar with regular and context-free languages, as well as some elementary properties of codes. For some background on this, and other topics in formal language theory, we refer the reader to standard books on the subject [23, 25]. The paper also assumes familiarity with the basics of the theory of monoid and group presentations, which will be written as Mon⟨A | R⟩ and Gp⟨A | R⟩, respectively. For further background information and examples of this theory, see e.g. [1, 31], [36, 30, 9].

We begin by fixing some notation and presenting some basics of rewriting systems. We refer the reader to the monographs by Jantzen [27] and Book & Otto [6] for a more thorough background on this topic, including its history.

Let A be a finite alphabet, and let A* denote the free monoid on A, with identity element denoted ε or 1, depending on the context. Let A+ denote the free semigroup on A, i.e. A+ = A \ {ε}. For u, v ∈ A+, by u ≡ v we mean that u and v are the same word. For w ∈ A*, we let |w| denote the length of w, inductively defined by |ε| = 0 and |wa| = |w| + 1 for w ∈ A* and a ∈ A. We use Pre(w) (resp. Suf(w)) to denote the set of non-empty prefixes (resp. non-empty suffixes) of the word w ∈ A*. If two distinct words u, v ∈ A+ are such that Pre(u) ∩ Suf(v) ≠ ∅, then we say that u and v overlap. We say that w ∈ A+ overlaps with itself if Pre(w) ∩ Suf(w) ≠ {w}, and we say that w self-overlap free (sometimes also called a bifix-free word or hypersimple word) if it does not overlap with itself. For example, the word xyzabcxyz is not self-overlap free, whereas arghmog, on the other hand, is.

Equality of words u, v ∈ A* in the monoid M = Mon⟨A | R⟩ is denoted u =_M v.

Let G be a group with finite (group) generating set A, with A−1 the set of inverses of the generators A. The language

\{ w \mid w ∈ (A ∪ A^{−1})^*, w =_G 1 \}

is called the (group-theoretic) word problem for G, see e.g. [33, 35]. Let M be a monoid with a finite generating set A. Translating the above definition of the word problem directly to M does not, in general, yield much insight into the structure of M. Duncan & Gilman [14, p. 522] instead introduced a different generalisation of the group-theoretic word problem to all monoids. The word problem of M with respect to A is defined as the language

\{u # v^{rev} \mid u, v ∈ A^*, u =_M v \},

where # is some symbol not in A, and v^{rev} denotes the reversal of v, i.e. the word v read backwards. We remark that although Duncan & Gilman initialised the proper study of this language-theoretic word problem for monoids, the language in the definition was studied already by Book, Jantzen & Wrathall in 1982, see especially [5, Corollary 3.8]. This observation does not seem to appear anywhere in the literature on the word problem for monoids. We say that M has context-free word problem if this language is context-free; this can be shown to not depend

\footnote{The charming word arghmog was invented, and used exactly once, by Turing in the introduction to his paper demonstrating the general undecidability of the word problem for finitely presented cancellative semigroups [49, p. 491].}
on the finite generating set chosen [13, Theorem 5.2]. Furthermore, a group has context-free word problem in the usual sense if and only if the above language is a context-free language [13, Theorem 3]. If context is clear, we will call a monoid context-free if its word problem is context-free. Note that non-finitely generated monoids are not context-free.

We now give the basic definitions regarding rewriting systems. A rewriting system \( R \) on \( A \) is a subset of \( A^* \times A^* \). An element of \( R \) is called a rule. The system \( R \) induces several relations on \( A^* \). We will write \( u \rightarrow_R v \) if and only if there exist \( x, y \in A^* \) and a rule \( (\ell, r) \in T \) such that \( u \equiv x\ell y \) and \( v \equiv xy \). We let \( \rightarrow^*_R \) denote the reflexive and transitive closure of \( \rightarrow_R \). We denote by \( \iff_R \) the symmetric and transitive closure of \( \rightarrow_R \). The relation \( \iff_R \) defines the least congruence on \( A^* \) containing \( \rightarrow_R^* \). The monoid \( \text{Mon}(A \mid R) \) will throughout this note, and without comment, be naturally be identified with the quotient \( A^*/\iff_R \).

If \( u \in A^* \) is such that there is no \( u \rightarrow_R v \), then \( u \) is said to be \( R \)-irreducible. We say that \( R \) is terminating (or Noetherian) if there is no infinite chain of the form \( u_1 \rightarrow_R u_2 \rightarrow_R \cdots \), and we say that it is confluent if for all \( u, v, w \in A^* \), we have that \( u \rightarrow_R v \) and \( u \rightarrow_R w \) together imply that there exists some \( z \in A^* \) with \( v \rightarrow_R z \) and \( w \rightarrow_R z \). We say that it is locally confluent if \( u \rightarrow_R v \) and \( u \rightarrow_R w \) together imply that there exists some \( z \in A^* \) with \( v \rightarrow_R z \) and \( w \rightarrow_R z \). Newman’s classical lemma (see [37]) says that a terminating and locally confluent rewriting system is confluent. If \( R \) is confluent and terminating, we say it is complete. A rewriting system \( R \subseteq A^* \times A^* \) is said to be monadic if \((u, v) \in R \) implies \(|u| > |v| \) and \( v \in A \cup \{\varepsilon\} \). We say that \( R \) is special if \((u, v) \in R \) implies \(|u| > |v| \) and \( v \equiv \varepsilon \). The monadic rewriting system \( R \) is said to be context-free (resp. regular) if for every \( a \in A \cup \{\varepsilon\} \), the set \( \{u \mid (u, a) \in R\} \) is a context-free (resp. regular) language.

We fix some terminology regarding monoids. Let \( M = \text{Mon}(A \mid R) \). An element \( m \in M \) is a left unit if there exists \( m' \in M \) such that \( m'm = 1 \) in \( M \). We define a right unit analogously. We say that \( m \) is a unit if it is both a left and right unit. The subgroup of \( M \) consisting of all units is called the group of units of \( M \), and is denoted \( U(M) \). We say that \( u \in A^* \) is left (resp. right) invertible in \( M \) if it represents a left (resp. right) unit in \( M \); and \( u \) is invertible in \( M \) if it represents a unit. For an idempotent \( e \in E(M) \) of \( M \), we define the maximal subgroup of \( M \) containing \( e \) to be the group of units of \( eMe \). The maximal subgroup of \( M \) containing \( e \) is a subsemigroup of \( M \). It is, additionally, a group – but note that its identity element will be \( e \), rather than \( 1 \). For an equivalent definition, the set of maximal subgroups of \( M \) is the set of all group \( \mathcal{H} \)-classes of \( M \) [11, Ch. 2]. If \( X \subseteq A^* \), we denote by \( \langle X \rangle_M \) the submonoid of \( M \) generated by the elements represented by the words in \( X \). Note that \( 1 \in \langle X \rangle_M \).

Finally, we say that a monoid of the form \( M = \text{Mon}(A \mid w_1 = 1 \ (i \in I)) \) is special, i.e. a monoid is special if the right-hand sides of all defining relations are empty. Such monoids were introduced by Tseitin [45], and were extensively studied by Adian [2] and Makanin [33, 32]. For a more in-depth analysis of special monoids and their history, we refer the reader to [38, 40, 42, 51]. The set \( \{w_i \mid i \in I\} \) is called the set of defining words of \( M \). We say that a non-empty invertible word is minimal if none of its non-empty proper prefixes is invertible in \( M \). We remark that minimality of \( w \) is clearly seen to be equivalent to \( w \) having none of its non-empty proper suffixes being invertible. We will use the term minimal invertible...
factor and minimal factor interchangeably. Each defining word is invertible in \( M \), and can hence be factored uniquely as \( w_i \equiv w_{i,1} w_{i,2} \cdots w_{i,\ell} \) into minimal factors. We set \( \Lambda(w_i) = \{ w_{i,1}, \ldots, w_{i,\ell} \} \), and let \( \Lambda = \Lambda(M) \) denote the set \( \bigcup_{w_i} \Lambda(w_i) \) of all minimal words arising from the factorisation of the defining words of \( M \). We say that \( \Lambda \) is the set of minimal words associated to \( M \). It is clear that \( \Lambda \) is a biprefix code, i.e. that \( \Lambda \cap A^* \Lambda = \Lambda \cap \Lambda A^* = \emptyset \).

3. Proof of the main theorem

Let \( \Pi_2 = \text{Mon}(a, b, c \mid (ab^i c)^2 = 1 \ (i \geq 1)) \). We will throughout this section let \( A = \{ a, b, c \} \). We will show that \( \Pi_2 \) is context-free, but that none of the maximal subgroups of \( \Pi_2 \) are finitely generated. We shall use the following lemma to provide a direct path to some simple structural results about the group of units of \( \Pi_2 \).

**Lemma 3.1.** The rewriting system \( \mathcal{T}_2 = \{((ab^i c)^2, e) \mid i \geq 1\} \subseteq A^* \times A^* \) is a context-free, complete rewriting system which defines \( \Pi_2 \).

**Proof.** Obviously, \( \mathcal{T}_2 \) defines \( \Pi_2 \). As \( \mathcal{T}_2 \) is length-reducing, it is terminating. To show that \( \mathcal{T}_2 \) is complete, it thus suffices to show that it is locally confluent; but this is immediate, as the only overlap between two left-hand sides of rules from \( \mathcal{T}_2 \) are of the form \( ab^i c(ab^j c)ab^k c \), where \( i \geq 1 \), and this overlap is immediately resolved. Hence \( \mathcal{T}_2 \) is complete. Finally, the set of left-hand sides of the empty word in \( \mathcal{T}_2 \) is the language \( \{ ab^i c(ab^j c) \mid i \geq 1 \} \), which is a context-free language. \( \square \)

As an immediate corollary of Lemma 3.1 we have the following.

**Proposition 3.2.** The special monoid \( \Pi_2 \) has context-free word problem.

**Proof.** Any monoid defined by a monadic complete rewriting system has context-free word problem [5, Corollary 3.8], and by Lemma 3.1 the system \( \mathcal{T}_2 \) is a special (and thus monadic) complete rewriting system defining \( \Pi_2 \). \( \square \)

**Lemma 3.3.** For all \( i \geq 1 \), the factorisation of the word \( (ab^i c)^2 \) into minimal invertible factors in \( \Pi_2 \) is \( (ab^i c)(ab^i c) \).

**Proof.** Let \( i \geq 1 \). As \( ab^i c \) is invertible in \( \Pi_2 \), any factorisation of \( (ab^i c)^2 \) into minimal invertible factors must obviously be a refinement of the factorisation \( (ab^i c)(ab^i c) \). We thus consider factorisations of \( ab^i c \). Suppose \( ab^i c \equiv \lambda_1 \lambda_2 \cdots \lambda_n \) where the \( \lambda_j \) (\( 1 \leq j \leq n \)) are minimal invertible factors. If \( n > 2 \), then there exists \( 1 < j < n \) such that \( \lambda_j \equiv b^{i \mu} \) for some \( 1 \leq \mu \leq i \). As \( \lambda_j \) is minimal, it is not a proper power of a word. Thus \( \mu = 1 \) and \( b \) is invertible. Hence, as \( \lambda_1 \equiv ab^\nu \) for some \( \nu \geq 0 \), it follows from the minimality of \( \lambda_1 \) that \( \nu = 0 \). Thus \( \lambda_1 \equiv a \) is invertible. But \( a \) is not invertible in \( \Pi_2 \); indeed, if it were, then it would be left invertible, so there would be some \( u \in A^* \) such that \( ua = 1 \) in \( \Pi_2 \). But no rule of the rewriting system \( \mathcal{T}_2 \) ends in \( a \), and hence, by induction of the number of rules applied, we cannot have \( ua \rightarrow^* \mathcal{T}_2 \). As \( \mathcal{T}_2 \) is complete and defines \( \Pi_2 \), we therefore have \( ua \neq 1 \) in \( \Pi_2 \), a contradiction.

Thus \( n \leq 2 \). If \( n = 2 \), then \( ab^i c \equiv \lambda_1 \lambda_2 \), and consequently \( \lambda_1 \equiv ab^\mu \) and \( \lambda_2 \equiv b^{i \nu} \) for some \( \mu, \nu \geq 1 \) with \( \mu + \nu = i \). But then \( b \), being a prefix of the invertible \( \lambda_2 \) and a suffix of the invertible \( \lambda_1 \), is both right and left invertible, and hence invertible. Thus \( \lambda_1 \equiv ab^\mu \) has a non-empty proper invertible suffix, contradicting its minimality. We conclude that \( n = 1 \), and thus the factorisation of \( ab^i c \) into minimal invertible factors is \( (ab^i c) \), whence the factorisation of \( (ab^i c)^2 \). \( \square \)
Let now $\Lambda = \{ab^ic \mid i \geq 1\}$. By Lemma 3.3 $\Lambda$ is the set of minimal words associated to $\Pi_2$. The following proof follows the proofs of [51] Lemmas 3.3 & 3.4 (which consider only finitely presented special monoids) but does not rely on the technical machinery preceding it; instead, we can directly use the rewriting system $T_2$.

**Lemma 3.4.** Let $u \in A^*$ be $T_2$-irreducible and invertible. Then $u \in \Lambda^*$.

**Proof.** We will first prove that if $w \in A^*$ is $T_2$-irreducible and right invertible, then $w \in \text{Pre}(\Lambda)^*$. The proof is by induction on $|w|$. If $|w| = 0$, then the claim is obvious. Suppose for induction that the claim is true for all right invertible words of length less than $|w|$, and that $|w| > 0$. As $w$ is right invertible, there exists some $w' \in A^*$ such that $ww' =_{\Pi_2} 1$. Assume without loss of generality that $w'$ is $T_2$-irreducible. As $T_2$ is complete and defines $\Pi_2$ by Lemma 3.1 we have $ww' \rightarrow_{T_2}^* 1$. As $w$ and $w'$ are $T_2$-irreducible, but $ww'$ is not as $|ww'| > 0$, we can hence write $w$ and $w'$ as $w \equiv w_0w_1$ and $w' \equiv w_0'w_1'$, with $|w_1| > 0$ and $|w_0'| > 0$, such that $ww' \equiv w_0w_1w_0'w_1' \rightarrow_{T_2} w_0w_1' \rightarrow_{T_2}^* 1$, and where $w_1w_0' \equiv (ab^ic)^2$ for some $i \geq 1$. As $w_1$ is a prefix of the word $w_1w_0' \equiv (ab^ic)^2 \in \Lambda^*$, we have $w_1 \in \text{Pre}(\Lambda)^*$. As $|w_1| > 0$, we have $|w_0| < |w|$. As $w_0$ is a prefix of the right invertible word and $T_2$-irreducible word $w$, it is itself right invertible and $T_2$-irreducible, and so by the inductive hypothesis $w_0 \in \text{Pre}(\Lambda)^*$. Thus $w \equiv w_0w_1 \in \text{Pre}(\Lambda)^*$, and we are done with the proof of the claim.

We are now ready to prove the lemma. Let now $u \in A^*$ be $T_2$-irreducible and invertible. Then $u$ is also right invertible; so, by the claim, let $p_1, p_2, \ldots, p_n \in \text{Pre}(\Lambda)$ be such that $u \equiv p_1p_2\cdots p_n$. As $p_1p_2\cdots p_n$ is invertible, it follows that $p_n$ is left invertible, being a suffix of an invertible word. But $p_n$ is right invertible, being a prefix of a word from $\Lambda$. Thus $p_n$ is invertible. But no non-empty proper prefix of a word from $\Lambda$ is invertible; hence $p_n \in \Lambda$. As $p_n$ and $p_1p_2\cdots p_{n-1}$ being invertible implies that $p_1p_2\cdots p_{n-1}$ is invertible, we can continue the same process to conclude that $p_1, p_2, \ldots, p_n \in \Lambda$. Thus $u \in \Lambda^*$.

The following proposition now follows directly from Lemma 3.1 and Lemma 3.4.

**Proposition 3.5.** The submonoid $\langle \Lambda \rangle_{\Pi_2}$ is the group of units of $\Pi_2$.

We note that, in the finitely presented case, the fact that the set of minimal invertible factors in general generates the group of units is [51] Lemma 3.4 or indeed [26] Prop 4.2; in the latter, it is proved for special *inverse* monoids via an algorithmic graphical procedure due to J. B. Stephen [45].

As no word in $\Lambda$ overlaps non-trivially with any other, we have that $\Lambda$ is a code. Using this, together with the fact that all defining words of $\Pi_2$ are elements of $\Lambda^*$, we can easily find a presentation for the submonoid $\langle \Lambda \rangle_{\Pi_2}$ of $\Pi_2$ using the presentation of $\Pi_2$. Malheiro [31] Theorem 2.4] proved a general presentation theorem for what he calls “codified” submonoids of monoids. We will only need a very special case of this, which we now state. For ease of notation, for the statement of the theorem we temporarily forget having defined $A = \{a, b, c\}$ and let $A$ be an arbitrary alphabet.

---

[4] The method, which will not be used in any capacity here but which has applications also to special monoids, is known as *Stephen’s procedure*, and was described already for groups in 1910 by Dehn [12] in his approximating construction of the *Gruppenbild*, or Cayley graph, of a group. For more details on Dehn’s construction, we refer the reader to [10] Chapter 1.5.
To every code $A_0 \subseteq A^+$ we associate an alphabet $X_0$ in bijective correspondence with $A_0$ via a map $\varphi: A_0 \to X_0$, which we extend to an isomorphism $\varphi: A_0^* \to X_0^*$.

**Theorem 3.6** (Malheiro [34]). Let $R = \{(u_1, \varepsilon),(u_2, \varepsilon), \ldots\}$ be a special complete rewriting system, on an alphabet $A$, which defines the monoid $M$. Let $A_0 \subseteq A^+$ be a code such that all left-hand sides of rules in $R$ are elements of $A_0^+$, and let $X_0, \varphi$ be as above. Then

$$\text{Mon}(X_0 \mid \varphi(u_1) = 1, \varphi(u_2) = 1, \ldots)$$

is a presentation for the submonoid $\langle A_0 \rangle_M$.

Redefining $A = \{a, b, c\}$, we can now apply the above theorem to our monoid.

**Proposition 3.7.** The group $U(\Pi_2)$ is isomorphic to the free product of infinitely many finite cyclic groups $\text{Gp}(x_i \mid i \geq 1 \mid x_i^2 = 1 (i \geq 1))$. In particular, $U(\Pi_2)$ is non-finitely generated.

**Proof.** As $\Lambda$ is a code, let $X = \{x_1, x_2, \ldots\}$ be an alphabet in bijective correspondence with $\Lambda$ via the map induced by $ab^ic \mapsto x_i$, and let $\varphi: \Lambda^* \to X^*$ be the homomorphism extending this map. As all left-hand sides of rules in $\mathcal{T}_2$ are words in $\Lambda^+$, it follows by Lemma 3.5 and Theorem 3.6 that

$$\langle \Lambda \rangle_{\Pi_2} \cong \text{Mon}(X \mid \varphi((ab^ic)^2) = 1)$$

$$\cong \text{Mon}(x_i \mid i \geq 1 \mid x_i^2 = 1) \cong \text{Gp}(x_i \mid i \geq 1 \mid x_i^2 = 1),$$

where the last isomorphism is due to the fact that the $x_i$ are all clearly invertible in the monoid $\text{Mon}(x_i \mid i \geq 1 \mid x_i^2 = 1)$. By Proposition 3.5 we have $\langle \Lambda \rangle_{\Pi_2} = U(\Pi_2)$, and the claim follows.

Thus we have proved the following.

**Theorem 3.8.** The special monoid $\Pi_2$ is a context-free monoid with non-finitely generated group of units.

As all maximal subgroups of a special monoid are isomorphic to its group of units [34, Theorem 4.6], we have thus proved the main theorem of this note.

**Corollary 3.9.** The special monoid $\Pi_2$ is a context-free monoid, none of whose maximal subgroups is finitely generated.

We remark that $U(\Pi_2)$ is not context-free (as it is not finitely generated), but it is locally context-free, i.e. every finitely generated subgroup of $U(\Pi_2)$ is context-free. This follows from the Kurosh subgroup theorem, and the fact that any finitely generated subgroup of a context-free group is itself context-free.

One can regard $\Pi_2$ as the result of a certain procedure: start with a family of pairwise isomorphic groups $C_{2,i} = \text{Mon}(x_i \mid x_i^2 = 1)$, each of which admits a context-free complete monadic rewriting system. Then, from a suitably chosen biprefix code (in this case $\{ab^ic \mid i \geq 1\}$), create a presentation for $\Pi_2$ by taking as generators $a, b, c$, and as defining relations the union over $i$ of all defining relations of $C_{2,i}$ subjected to the substitution $x_i \mapsto ab^ic$. Of course, more generally, we can define $\Pi_n = \text{Mon}(a, b, c \mid (ab^ic)^n = 1 (i \geq 1))$ in the same manner. It is not hard to show that the earlier analysis carries over, mutatis mutandis, to show that $U(\Pi_n)$ is a free product of infinitely many finite cyclic groups of order $n$.

However, unlike in the case $n = 2$, for $n > 3$ the language $\{(ab^ic)^n \mid i \geq 1\}$ is not context-free; this can be proved using the pumping lemma for context-free
languages, and mirrors the fact that \( \{ x_i x_j \cdots x_{i}^{n} \mid i \geq 1 \} \) is a context-free language if and only if \( n \leq 2 \). Indeed, using a rewriting system \( T_n \) defined in the same manner as \( T_2 \), it is clear that \( WP_{\Pi_n} \cap (ab^c)^n \# \) is the reversal of the language \( \{ (ab^c)^n \mid i \geq 1 \} \), and hence \( WP_{\Pi_n} \) is not context-free when \( n > 2 \), as the context-free languages are closed under intersection with regular languages and reversal. Thus the special monoid \( \Pi_n \) has context-free word problem if and only if \( n \leq 2 \).

The above construction is quite general. We provide one potential broader usage of it. Let \( G_n \) be the elementary abelian 2-group \( C_2^m \) of order \( m = 2^n \), and let \( x_1, \ldots, x_{m-1} \) be its non-trivial elements. Then \( G_n \) admits a finite presentation via the multiplication table of \( G_n \) as

\[
G_n = \text{Gp}\langle x_1, \ldots, x_{m-1} \mid R_n \rangle = \text{Mon}\langle x_1, \ldots, x_{m-1} \mid R_n \rangle
\]

where the relations \( R_n \) are all of the form \( x_i x_j = x_k \) or \( x_i x_j = 1 \) where \( (1 \leq i \leq j \leq k \leq m-1) \). Note that the equality between the group resp. the monoid presentation is due to the presence of relations of the form \( x_i^2 = 1 \) for all \( 1 \leq i \leq m-1 \), so every \( x_i \) is invertible in the monoid presentation. Let \( b_1, \ldots, b_{m-1} \) be new symbols. Analogously to how \( \Pi_n \) was obtained from \( C_n \), we define a new monoid \( M_n \) by, for every \( i \geq 1 \), rewriting the relations in \( R_n \) to a set of relations \( R_{i,n} \) over the alphabet \( \{ a, b_1, b_2, \ldots, b_{m-1}, c \} \) by performing the substitution \( x_j \mapsto ab_j^i c \), and then defining

\[
M_n = \text{Mon}\langle a, b_1, \ldots, b_n, c \mid \bigcup_{i \geq 1} R'_{i,n} \rangle.
\]

One may show, in the same manner as before, that the rewriting system \( \bigcup_{i \geq 1} R'_{i,n} \) is complete (as \( R_n \) is complete), and that \( U(M_n) \) is a free product of countably many copies of the elementary abelian 2-group \( G_n \), and hence is not context-free. While the rewriting system is not monadic when \( n > 2 \), its behaviour appears to the author as rather similar to that of a context-free monadic rewriting system. We therefore suspect that \( M_n \) has context-free word problem, but do not have a proof.

We end this note with a connection to coherent groups. That the group of units of \( \Pi_2 \) is a free product of countably many finitely generated virtually free groups has many structural implications. Indeed, it follows directly from the Kurosh subgroup theorem that, in an infinite free product of finitely presented virtually free groups, any finitely generated subgroup is finitely presented. In particular, \( U(\Pi_2) \) is a coherent group (a group \( G \) is coherent if every finitely generated subgroup of \( G \) is also finitely presented). Furthermore, \( U(\Pi_2) \) is locally free (resp. locally context-free), i.e. every finitely generated subgroup of \( U(\Pi_2) \) is virtually free (resp. context-free). This raises the following question of how bad the structure of the group of units of a context-free monoid can be.

**Question 3.10.** Does there exist a context-free monoid with an incoherent maximal subgroup? Does there exist such a monoid which is special?

This question can also be posed with incoherent replaced by not locally context-free. Coherence of groups, particularly that of one-relator groups, has recently seen a good deal of study (see especially [50],[1],[29]), and recent work by Gray & Ruškuc [22] has linked the coherence of one-relator groups to the question of groups of units of special inverse monoids. This makes the question of coherence particularly interesting. As any finitely presented context-free special monoid necessarily has context-free group of units (see [40]), any special monoid witnessing a positive answer to the question is necessarily non-finitely presentable.
References

[1] S. I. Adian, Defining relations and algorithmic problems for groups and semigroups, Proceedings of the Steklov Institute of Mathematics, No. 85 (1966).

[2] The problem of identity in associative systems of a special form, Soviet Math. Dokl. 1 (1960), 1360–1363, [Dokl. Akad. Nauk SSSR 135, 6 (1960)].

[3] A. V. Anisiimov, The group languages, Kibernetika (Kiev) (1971), no. 4, 18–24.

[4] Gilbert Baumslag, Benjamin Fine, and Gerhard Rosenberger, One-relator groups: an overview, Groups St Andrews 2017 in Birmingham, London Math. Soc. Lecture Note Ser., vol. 455, Cambridge Univ. Press, Cambridge, 2019, pp. 119–157.

[5] Ronald V. Book, Matthias Jantzen, and Celia Wrathall, Monadic Thue systems, Theor. Comput. Sci. 19 (1982), no. 3, 231–251.

[6] Ronald V. Book and Friedrich Otto, String-rewriting systems, Texts and Monographs in Computer Science, Springer-Verlag, New York, 1993.

[7] Tara Brough, Groups with poly-context-free word problem, Groups Complex. Cryptol. 6 (2014), no. 1, 9–29.

[8] Tara Brough, Alan J. Cain, and Markus Pfeiffer, Context-free word problem semigroups, Developments in language theory, Lecture Notes in Comput. Sci., vol. 11647, Springer, Cham, 2019, pp. 292–305.

[9] C. M. Campbell, E. F. Robertson, N. Ruškuc, and R. M. Thomas, Semigroup and group presentations, Bull. London Math. Soc. 27 (1995), no. 1, 46–50. MR 1331680

[10] Bruce Chandler and Wilhelm Magnus, The history of combinatorial group theory, Studies in the History of Mathematics and Physical Sciences, vol. 9, Springer-Verlag, New York, 1982, A case study in the history of ideas. MR 680777

[11] A. H. Clifford and G. B. Preston, The algebraic theory of semigroups. Vol. I, Mathematical Surveys, No. 7, American Mathematical Society, Providence, R.I., 1961.

[12] M. Dehn, Über die Topologie des dreidimensionalen Raumes, Math. Ann. 69 (1910), no. 1, 137–168.

[13] Andrew Duncan and Robert H. Gilman, Word hyperbolic semigroups, Math. Proc. Cambridge Philos. Soc. 136 (2004), no. 3, 513–524.

[14] J. Duncan and A. L. T. Paterson, Amenability for discrete convolution semigroup algebras, Math. Scand. 66 (1990), no. 1, 141–146.

[15] M. J. Dunwoody, The accessibility of finitely presented groups, Invent. Math. 81 (1985), no. 3, 449–457.

[16] Robert H. Gilman, Robert P. Kropholler, and Saul Schleimer, Groups whose word problems are not semilinear, Groups Complex. Cryptol. 10 (2018), no. 2, 53–62.

[17] É. A. Golubov, Finitely approximable regular semigroups, Mat. Zametki 17 (1975), 423–432.

[18] R. Gray, A. Malheiro, and S. J. Pride, On properties not inherited by monoids from their Schützenberger groups, Inform. and Comput. 209 (2011), no. 7, 1120–1134.

[19] R. Gray and S. J. Pride, Homological finiteness properties of monoids, their ideals and maximal subgroups, J. Pure Appl. Algebra 215 (2011), no. 12, 3005–3024.

[20] R. Gray and N. Ruškuc, Green index and finiteness conditions for semigroups, J. Algebra 320 (2008), no. 8, 3145–3164.

[21] , Generators and relations for subsemigroups via boundaries in Cayley graphs, J. Pure Appl. Algebra 215 (2011), no. 11, 2761–2779.

[22] Robert D. Gray and Nik Ruškuc, On groups of units of special and one-relator inverse monoids, Pre-print (2021), Available at arXiv:2103.02995.

[23] Michael A. Harrison, Introduction to formal language theory, Addison-Wesley Publishing Co., Reading, Mass., 1978.

[24] Derek F. Holt, Sarah Rees, Claas E. Röver, and Richard M. Thomas, Groups with context-free co-word problem, J. London Math. Soc. (2) 71 (2005), no. 3, 643–657.

[25] John E. Hopcroft and Jeffrey D. Ullman, Introduction to automata theory, languages, and computation, Addison-Wesley Publishing Co., Reading, Mass., 1979, Addison-Wesley Series in Computer Science.

[26] S. V. Ivanov, S. W. Margolis, and J. C. Meakin, On one-relator inverse monoids and one-relator groups, J. Pure Appl. Algebra 159 (2001), no. 1, 83–111.

[27] Matthias Jantzen, Confluent string rewriting, EATCS Monographs on Theoretical Computer Science, vol. 14, Springer-Verlag, Berlin, 1988.
[28] Andrzej Jura, *Determining ideals of a given finite index in a finitely presented semigroup*, Demonstratio Math. **11** (1978), no. 3, 813–827.

[29] Larsen Louder and Henry Wilton, *One-relator groups with torsion are coherent*, Math. Res. Lett. **27** (2020), no. 5, 1499–1512.

[30] Roger C. Lyndon and Paul E. Schupp, *Combinatorial group theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 89, Springer-Verlag, Berlin-New York, 1977.

[31] Wilhelm Magnus, Abraham Karrass, and Donald Solitar, *Combinatorial group theory: Presentations of groups in terms of generators and relations*, Interscience Publishers [John Wiley & Sons, Inc.], New York-London-Sydney, 1966.

[32] G. S. Makanin, *On the Identity Problem for Finitely Presented Groups and Semigroups*, Ph.D. thesis, Steklov Mathematical Institute, Moscow, 1966, English translation by C.-F. Nyberg-Brodda available at arXiv:2102.00745.

[33] On the identity problem in finitely defined semigroups, Soviet Math. Dokl. **7** (1966), 1478–1480, [Dokl. Akad. Nauk SSSR **171** (1966)].

[34] António Malheiro, *Complete rewriting systems for codified submonoids*, Internat. J. Algebra Comput. **15** (2005), no. 2, 207–216.

[35] David E. Muller and Paul E. Schupp, *Groups, the theory of ends, and context-free languages*, J. Comput. System Sci. **26** (1983), no. 3, 295–310.

[36] B. H. Neumann, *Some remarks on semigroup presentations*, Canadian J. Math. **19** (1967), 1018–1026.

[37] M. H. A. Newman, *On theories with a combinatorial definition of “equivalence.”*, Ann. of Math. (2) **43** (1942), 223–243.

[38] C.-F. Nyberg-Brodda, The geometry of special monoids, Pre-print (submitted) (2020), Available at arXiv:2011.04536.

[39] On the word problem for compressible monoids, Pre-print (submitted) (2020), Available at arXiv:2012.01402.

[40] On the word problem for special monoids, Pre-print (submitted) (2020), Available at arXiv:2011.09466.

[41] The Word Problem and Combinatorial Methods for Groups and Semigroups, Ph.D. thesis, University of East Anglia, UK, 2021, In preparation.

[42] The word problem for one-relation monoids: a survey, Semigroup Forum **103** (2021), To appear.

[43] N. Ruskuc, *Presentations for subgroups of monoids*, J. Algebra **220** (1999), no. 1, 365–380.

[44] N. Ruskuc and R. M. Thomas, *Syntactic and Rees indices of subsemigroups*, J. Algebra **205** (1998), no. 2, 435–450.

[45] J. B. Stephen, *Applications of automata theory to presentations of monoids and inverse monoids*, ProQuest LLC, Ann Arbor, MI, 1987, Thesis (Ph.D.)–The University of Nebraska - Lincoln.

[46] A. K. Sushkevich, *The theory of operations as the general theory of groups*, Ph.D. thesis, Voronezh State University, 1922.

[47] Axel Thue, *Probleme über Veränderungen von Zeichenreihennach gegebenen Regeln*, Christiana Videnskabs-Selskabs Skrifter **10** (1914).

[48] G. S. Tseitin, *An associative calculus with an insoluble problem of equivalence*, Trudy Mat. Inst. Steklov. **52** (1958), 172–189.

[49] A. M. Turing, *The word problem in semi-groups with cancellation*, Ann. of Math. (2) **52** (1950), 491–505.

[50] Daniel T. Wise, *The coherence of one-relator groups with torsion and the Hanna Neumann conjecture*, Bull. London Math. Soc. **37** (2005), no. 5, 697–705.

[51] Louxin Zhang, *Applying rewriting methods to special monoids*, Math. Proc. Cambridge Philos. Soc. **112** (1992), no. 3, 495–505.

**DEPARTMENT OF MATHEMATICS, ALAN TURING BUILDING, UNIVERSITY OF MANCHESTER, UK.**

**Email address:** carl-fredrik.nybergbrodda@manchester.ac.uk