Finiteness properties for semigroups and their substructures

A thesis submitted to the School of Mathematics at the University of East Anglia in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Diane Christine Lubbock

March 2021

Abstract

In this thesis we consider finiteness properties of infinite semigroups and infinite monoids. In particular we investigate finite presentations which have the property finite derivation type (FDT) or the property that they admit a presentation by a finite complete rewriting system (FCRS). We ask the question of whether these properties are inherited between a semigroup (or monoid) and particular substructures like subsemigroups (or submonoids).

We first investigate completely simple semigroups (which are isomorphic to Rees matrix semigroups) that have a single $\mathcal{R}$-class or a single $\mathcal{L}$-class. We prove that the maximal subgroups admit a presentation by a FCRS if and only if the semigroup admits a presentation by a FCRS with respect to a sparse generating set. Next we move on to our second stream of research and consider the property that a presentation has FDT. We study unitary subsemigroups with finite strict boundary (a condition given in terms of the Cayley graph) and prove that such subsemigroups inherit the property of FDT.

We prove that every finitely generated subsemigroup of the Bicyclic monoid admits a presentation by a FCRS. Finally we investigate FDT and FCRS for finitely generated submonoids of Plactic monoids, proving that these properties are satisfied in several cases. We make use of the fact that the Plactic monoid is known for having elements which correspond to semistandard tableau.

Contents

List of Figures

Acknowledgements

I would like to thank my supervisor Dr Robert Gray for his guidance and endless patience throughout my PhD. Also to thank my second supervisor Prof. Shaun Stevens for his support and for giving me the opportunity to study at the UEA. I would like to thank the university staff and my fellow students for all their help and advice. I would particularly like to thank Dr Robert Penfold who has encouraged me in my academic journey from my first beginnings with the Open University.

I would like to thank my friends Faith, Taka, Karen and Robert for their ceaseless support and encouragement through some difficult times and for reminding me that there is life outside of a PhD. Also to thank Jodie for always listening to my worries and for her common sense advice as someone who completed her PhD journey before me.

Finally, I would like to especially thank my long suffering partner, Chris, without whose support this would not have been possible. In particular, for his constant encouragement during the Covid-19 lockdown and my final months of writing this thesis.

Dedications

To my late Mother and Father.

1

Introduction

1.1 Related research and motivation

The areas of research included in this document relate to general finiteness properties of infinite semigroups. A finiteness property is one which holds for all finite semigroups and so we look to see where this is also true for an infinite semigroup. So for example, being finite is a finiteness property, since every finite semigroup has this property. Other important finiteness conditions include the properties of being finitely generated or finitely presented. It is a fact that not all infinite semigroups are finitely presented or indeed generated. Research has been carried out to determine whether these two properties are shared between a semigroup and substructures of the semigroup, often for particular types of semigroup. Of interest is whether a property is inherited by a substructure and whether it is passed up from the substructure to the parent semigroup.

In general if S is a semigroup and T is a subsemigroup of S then the finiteness properties that hold in S will not necessarily be inherited by T. Similarly T may satisfy finiteness properties that S does not. So it is natural to ask under what conditions will properties be passed from S to T or vice versa. For example, if S is a group and T is a subgroup of finite index, then there are many interesting theorems in the literature which show that S and T must then share many finiteness properties. For example, the properties of finiteness, being finitely generated, finite presentability, having a soluble word problem, periodicity, local finiteness, and residual finiteness are all known to be preserved by taking finite index subgroups and under taking finite index extensions, we refer the reader to [35]. For semigroups the analogous results have been proved in the case that $S \setminus T$ is finite (so called, large subsemigroups), see the following papers for some examples [48] [53] [54].

In general, more results have been proved by showing that properties are passed up to the semigroup in certain situations. It is generally accepted that proofs in the opposite direction passing from S to T are much harder. For example, it is an open question for groups whether the property of admitting a finite complete presentation is preserved under taking finite index subgroups, see [44], and it is an open question whether the property of finite derivation type is inherited by large subsemigroups, see [40].

Much work has been carried out by N. Ruškuc for example [48] on large subsemigroups where it is proved that if T is a large subsemigroup of S (i.e. $S \setminus T$ is finite), then S is finitely generated (respectively presented) if and only if T is. In the same paper this is also proved for the property of being residually finite. In another paper [46] he proves that a regular monoid with finitely many left and right ideals is finitely presented if and only if all its maximal subgroups are finitely presented.

Rees matrix semigroups were introduced by D. Rees in his paper [45], published in 1940. They were also implicit in a paper [52] by A. Suschkewitsch published in 1928. Rees matrix semigroups have become a widely used semigroup construction with many applications relating to regular semigroups, see [41] for a survey. We will make use of the Rees theorem which states that a semigroup is completely simple if and only if it is isomorphic to a Rees matrix semigroup $\mathcal{M}[A; I, \Lambda; P]$ where A is a group. J.M. Howie and N. Ruškuc looked at constructions and presentations for monoids in [26] and in one result they derive a presentation for the Rees matrix semigroup $S = \mathcal{M}^0[A; I, \Lambda; P]$ where A is a monoid. H. Ayik and N. Ruškuc looked at generators and relations of Rees matrix semigroups in

[1]. They prove that under certain extra (finiteness) conditions a Rees matrix semigroup $\mathcal{M}[S; I, J; P]$ is finitely generated (finitely presented) if and only if S is finitely generated (finitely presented). Of particular interest is a recent paper [40] by A. Malheiro where he proves that a finitely presented completely simple semigroup $\mathcal{M}[G; I, \Lambda; P]$ has finite derivation type (FDT) if and only if the group G has finite derivation type.

C.M. Campbell, E.F. Robertson, N. Ruškuc and R.M. Thomas worked together on several papers including one on subsemigroups of finitely presented semigroups in [6]. This paper contains a summary of results and problems open at that time for both general semigroups and free semigroups with respect to various finiteness properties (finite generation, finite presentation, finite index) and substructures (subsemigroups, ideals and one-sided ideals). They include useful examples of where such properties are not inherited. In another paper [7] the same collaboration look at finite presentation and ideals, proving that an ideal of a finitely presented semigroup is not necessarily finitely presented, even if it is finitely generated as a semigroup.

A key paper when considering the properties finite generation and finite presentation is [5], titled Reidemeister-Scheier type rewriting for semigroups by C.M. Campbell, E.F. Robertson, N. Ruškuc and R.M. Thomas. Given a finitely generated and presented semigroup, this method determines a presentation for a subsemigroup that is generated by a given finite set of generators. The resulting presentation is an infinite one but under some circumstances it can be bounded to a finite set. This method is a semigroup analogue of that by Reidemeister-Shreier which is for groups, see [36] for further details.

More recently R.D. Gray and N. Ruškuc looked specifically at generators and relations for subsemigroups via boundaries in Cayley graphs [21]. The main result they prove is that given a finitely presented semigroup S and a subsemigroup T of S with T having a finite boundary in S, then T is finitely presented. Here the boundary of T in S is a certain subset of T obtained by looking at directed edges in the Cayley graph of S that begin outside T and terminate in T. This result generalises the corresponding theorems for large subsemigroups from [48]. A related result is given in [16] where R.D. Gray considered a specific form of subsemigroup, namely where a semigroup is left (respectively right) unitary with strict right (respectively left) boundary. He proves that such subsemigroups are finitely generated (respectively finitely presented) if the semigroup is finitely generated (respectively finitely presented).

Research in this thesis looks at the property of a presentation for a semigroup being a finite complete presentation also known as a finite complete rewriting system (FCRS) and whether this property is inherited by its substructures, for example subsemigroups. The property of a presentation being a FCRS is of interest as it identifies semigroups where the word problem can be solved. Briefly, the word problem is as follows. Let S be a semigroup generated by a finite set A. The word problem asks whether there is an algorithm which takes any two words u, v over A and decides whether or not $u = v$ in S. In general the word problem is not decidable, even for finitely presented semigroups. If, however, a semigroup admits a presentation by a finite complete rewriting system, then it may be shown that the word problem for the semigroup is decidable. In terms of words within a semigroup, we are looking to see if they can be rewritten to the same word under the given rewrite rules. This property is of interest in the development of theoretical computer languages and algorithms where it has played a major role in their development, see [3, Introduction]. The property of being complete (also known as convergent) leads to the idea of "effective computability" and the ability to solve the word problem in linear time.

In a recent paper [54], K.B. Wong and P.C. Wong prove that the property FCRS is passed down from a semigroup S to a subsemigroup T, providing $S \setminus T$ is finite i.e. T is a large subsemigroup of S. As part of their proof they establish a method for finding a presentation for the subsemigroup and providing certain properties hold, then the presentation is complete. Their result complements work by J. Wang [53] which proves inheritance in the opposite direction when $S \setminus T$ is finite.

Recently, homotopical methods have been used to study semigroups. In particular, C.C. Squier developed a homotopy theory for monoids. He constructed a *derivation graph* from a presentation for a semigroup and defined equivalence relations on the closed paths which he called *homotopy relations* and a set of closed paths called a *homotopy base.* A presentation is said to have finite derivation type (FDT) if it has a finite homotopy base. In fact, he showed that FDT is a property which is true for a semigroup, irrespective of the particular presentation. See Chapter 4 below for formal definitions of these notions.

More recent work [51] by C.C. Squier, F. Otto and Y.Kobayashi on finiteness conditions for rewriting systems, looks at the link between FRCS and FDT. They show that if a presentation for a semigroup is a FCRS then that implies the semigroup has FDT. Thus, proving or disproving that a semigroup has FDT can be a useful step on the way to considering whether it has a FCRS. In addition Y. Kobayashi shows in his paper [29] that every one-relator monoid has FDT; it is not known whether they also have a FCRS. A useful survey can be found in [43] by F. Otto and Y. Kobayashi which contains more background information and a summary with respect to the properties of FCRS, finite derivation type (FDT) and homological finiteness conditions (FP_n).

More recently R.D. Gray and A. Malheiro considered finite derivation type and inheritance between the semigroup and its subgroups, see [19]. They proved that if S is a regular monoid with finitely many left and right ideals, then S has FDT if and only if every maximal subgroup of S has FDT. In related work, in [18] it is shown that if S is a regular semigroup with finitely many left and right ideals, and if all maximal subgroups of S have a FCRS, then S also has a FCRS. The converse remains an open problem and is part of the motivation for the work we do in Chapter 3 below.

The first published description of the bicyclic monoid was given by E. Lyapin in 1953. A. H. Clifford and G. Preston claim that whilst working with D. Rees it was independently discovered at some point before 1943. In [10] it is referred to as the simplest member of an extensive class of semigroups known as the bisimple inverse semigroups with identity element. As such it is a very useful semigroup in the theory of simple semigroups. Two papers [13] and [14] by L. Descalço and N. Ruškuc prove some interesting properties of subsemigroups of the bicyclic monoid. In [13] they prove that any subsemigroup of the bicyclic monoid falls into one of five different forms. They go on to prove in [14] that all finitely generated subsemigroups are finitely presented. These two papers motivate the research into finitely presented subsemigroups of the bicyclic monoid that we undertake in this thesis, see Chapter 5 below.

The plactic monoid originated from work by C. Schensted [49] in 1961 and by D.E. Knuth [28] in 1970. It has been used in connection with problems in representation theory, algebraic combinatorics and theory of quantum groups. We refer the reader to [15] [31] [33] [32] [8] for further information. Various papers have proved that the plactic monoid has alternative presentations which also have the property of being a FCRS. In [4] A.J. Cain, R.D. Gray and A. Malheiro prove that for all $n \geq 1$ the plactic monoid P_n admits a finite complete presentation. The complete presentation they obtain is given with respect to a particular finite generating set which they call column generators. In earlier work [30] Kubat and Okniński use the Knuth-Bendix completion procedure [3, Section 2.4] to prove that the plactic monoid P_3 admits a finite complete presentation with respect to the natural degree-lexicographic ordering over the usual generating set $\{1, 2, 3\}$. They also prove that the same approach does not work for P_n where n is greater than three. These papers motivate the work we do below in Chapter 6 where we investigate the properties FCRS and FDT for submonoids of plactic monoids.

1.2 Summary of results

Research in this thesis broadly divides into two streams. The first concerns finite complete rewriting systems (FCRS) where the main focus has been on investigating conditions under which this property is passed from a semigroup to a subsemigroup and vice versa. The second stream concerns analogous questions for a related finiteness property called finite derivation type (FDT).

Chapter 2 contains basic definitions, including those for semigroups, presentations and rewriting systems. Various useful lemmas, theorems and techniques are also stated for reference in subsequent chapters. Where definitions are specific to a single chapter they are stated within that chapter, rather than in Chapter 2.

In Chapter 3 we consider inheritance of the property of being a FCRS by maximal subgroups of certain forms of Rees matrix semigroups. Given a semigroup with finitely many left and right ideals we would like to prove that FCRS is inherited by the maximal subgroups of the semigroup. We investigate completely simple semigroups and the special case where the semigroup S has a single $\mathcal{R}$-class or a single $\mathcal{L}$-class and prove results with respect to a specific form of generating set for S called a sparse generating set (see Definition 3.3.2). The main result in Chapter 3 is:

Theorem 1.2.1. *Let $S = \mathcal{M}[G; I, \Lambda; P]$ be a Rees matrix semigroup where G is a group and one of two conditions is true:*
(i) $I = \{1\}$ and $\Lambda = \{1, 2, \ldots, n\}$;
(ii) $I = \{1, \ldots, m\}$ and $\Lambda = \{1\}$.

Then the group G admits a presentation by a finite complete rewriting system if and only if S admits a presentation by a finite complete rewriting system with respect to some sparse generating set.

In Chapter 4 we investigate properties of a certain type of subsemigroup, in particular, unitary subsemigroups with finite strict boundaries. We begin with some basic definitions relating to various forms of boundaries. Then we define specific properties where a subsemigroup is left (respectively right) unitary with finite strict right (respectively left) boundary in S. Given this property, we look to build on an existing result in [16] by R.D. Gray which proves that such subsemigroups are finitely presented, providing the semigroup is finitely presented. Interestingly, there is a link with Chapter 3 as the subgroups in Theorem 1.2.1 part (i) above, are in fact left unitary with finite strict right boundary, with part (ii) being the dual.

Motivated by the above results we consider the property of finite derivation type (FDT), which has close connections to FCRS and forms the second stream of research in this thesis. Definitions and descriptions for FDT are included in Chapter 4. We consider the property of FDT being inherited by left unitary subsemigroups with finite strict right boundary in S. The main result in Chapter 4 is:

Theorem 1.2.2. *Let S be a finitely presented semigroup with T a subsemigroup of S. Suppose S has finite derivation type. Then:*
(i) if T is left unitary and has finite strict right boundary in S then T also has finite derivation type;
(ii) if T is right unitary and has finite strict left boundary in S then T also has finite derivation type.

In Chapter 5 we return to the property of a presentation being a FCRS and consider a widely studied monoid, namely the bicyclic monoid. The bicyclic monoid admits a presentation which is a FCRS and so we consider substructures of this monoid with respect to the inheritance of being a FCRS. To this end, of particular interest are the results obtained in the two papers [13] and [14] by L. Descalço and N. Ruškuc where they prove some interesting properties of subsemigroups of the bicyclic monoid. This chapter includes an introduction to the bicyclic monoid and the necessary results from these two papers. Building on this work we prove the main result in Chapter 5:

Theorem 1.2.3. *Let* $\mathbf{B}$*, defined by the presentation* $\langle b, c \mid bc = 1\rangle$*, be the bicyclic monoid. Then every finitely generated subsemigroup of* $\mathbf{B}$ *admits a presentation by a finite complete semigroup rewriting system.*

In Chapter 6 we consider the inheritance of the property of being a FCRS, this time with respect to substructures of the plactic monoid (referred to as P_n). As with the bicyclic monoid, the plactic monoid admits a finite complete presentation. Definitions and a classic presentation for the plactic monoid are included, together with an introduction to the construct called a Young tableau which proves useful when working with this monoid. In contrast to the bicyclic monoid, there is currently no classification of the submonoids of the plactic monoid. Therefore, research in this chapter considers various submonoids as generated by generating sets of specific forms. Initial results in this chapter relate to the isomorphism of certain submonoids with the plactic monoid and with other specific submonoids, as follows:

Theorem 1.2.4. *Let* P_n *be the plactic monoid generated by* $A = \{1, 2, \ldots, n\}$ *where* $n \in \mathbb{N}$*. Let* S *be the submonoid of* P_n *generated by* $A_q = \{1^q, 2^q, \ldots, n^q\}$ *for some fixed* $q \in \mathbb{N}$*. Then* S *is isomorphic to* P_n*.*

Theorem 1.2.5. *Let* P_n *be the plactic monoid generated by* $A = \{1, 2, \ldots, n\}$ *where* $n \in \mathbb{N}$*. Let* S *be the submonoid of* P_n *generated by* $A_s = \{1^{s_1}, 2^{s_2}, \ldots, n^{s_n}\}$ *for* $s_1, s_2, \ldots, s_n \in \mathbb{N}$ *and set* $q = gcd\ \{s_1, s_2, \ldots, s_n\}$*. Let* T *be the submonoid generated by* $A_t = \{1^{t_1}, 2^{t_2}, \ldots, n^{t_n}\}$ *where* $t_x = s_x/q$ *for all* $x \in \{1, 2, \ldots, n\}$*. Then* S *is isomorphic to* T*.*

In the next part of Chapter 6 we consider the more general case of monoids with a homogeneous presentation and certain forms of submonoid, again with respect to being a FCRS. A presentation is homogeneous if for every defining relation $u = v$, the words u and v have the same length as each other. This work is motivated by the fact that the plactic monoid has a homogeneous presentation. This research uses a result in the paper [37] by A. Malheiro where certain monoids can be generated by a code (which is a subset of the free monoid), and if they conform to specific conditions, then a presentation can be constructed which is a finite complete rewriting system. The results proved in this section include the main result plus a couple of applications:

Theorem 1.2.6. *Let M be the monoid defined by the homogeneous presentation $\langle A \mid R\rangle$ where A is a finite alphabet and R is a finite set of relations. Fix $j \in \mathbb{N}$ and let E be the submonoid of M which consists of all elements of M which have a length divisible by j.*

Suppose the presentation $\langle A \mid R\rangle$ is a finite complete rewriting system. Then there exists a presentation for E which is a finite complete rewriting system.

Corollary 1.2.7. *Let P_3 be the plactic monoid of rank 3 defined by the ordered alphabet $A = \{1, 2, 3\}$ and let*

$$R = \{(332, 323), (322, 232), (331, 313), (311, 131), (221, 212), (211, 121),$$

$$(231, 213), (312, 132), (3212, 2321), (32131, 31321), (32321, 32132)\}.$$

Then P_3 is defined by the presentation $\langle A \mid R\rangle$. Fix $j \in \mathbb{N}$ and let E_j be the submonoid of P_3 containing only elements of P_3 of length divisible by j.

Then there exists a presentation for E_j which is a finite complete rewriting system.

We also prove the following result which generalises Theorem 1.2.6.

Theorem 1.2.8. *Let M be a monoid defined by a finite homogeneous presentation $\langle A \mid R\rangle$. Let $\psi : M \to (\mathbb{N}, +)$ be the surjective homomorphism induced by the mapping $w \mapsto |w|$ for $w \in A^*$. Let T be a finitely generated submonoid of $(\mathbb{N}, +)$. Then*

(i) $N = T\psi^{-1}$ is a finitely presented submonoid of M.

(ii) Moreover, if (A, R) is a finite complete rewriting system then $N = T\psi^{-1}$ admits a presentation by a finite complete rewriting system.

The remainder of Chapter 6 focuses on the plactic monoid P_2 and submonids generated by sets of the form $\{1, 2^i\}$ and $\{1^i, 2\}$ where $i \in \mathbb{N}$. Again we investigate whether these submonoids admit a presentation which is a finite complete rewriting system. Note that a previous result in this chapter, namely Theorem 1.2.5, can then be applied to widen the application of the following theorems to include generating sets of the form $\{1^q, 2^{qi}\}$ and $\{1^{qi}, 2^q\}$. The main result is:

Theorem 1.2.9. *Let $A = \{1, 2\}$ and $R = \{(221, 212), (211, 121)\}$. Then the plactic monoid P_2 is defined by the presentation $\langle A \mid R\rangle$.*

Fix $i \in \mathbb{N}$ with $i \geq 1$ and let E be the submonoid of P_2 generated by $X = \{1, 2^i\}$. Set $B = \{a, b\}$ and let Q be the subset of $B^* \times B^*$ consisting of all the following pairs:
(i) (bba, bab),
(ii) (ba^{i+1}, aba^i),
(iii) $\{(ba^{j-1}ba, ba^jb) \ : \ 2 \leq j \leq i\}$.
Then (B, Q) is a finite complete rewriting system defining the monoid E where a and b correspond to the generators 1 and 2^i, respectively.

Finally in Chapter 6, we look to generalise these results to submonoids generated by the set $\{1^i, 2^j\}$ where $i < j$ and they are co-prime. We obtain a result for the specific generating set $\{1^2, 2^3\}$, which is included below. A partial conjecture for the submonoid generated by $\{1^i, 2^j\}$ can be found at the end of Chapter 6.

Theorem 1.2.10. *Let $A = \{1, 2\}$ and $R = \{(221, 212), (211, 121)\}$. Then the plactic monoid P_2 is defined by the presentation $\langle A \mid R\rangle$.*

Let F_1 be the submonoid of P_2 generated by $Y_1 = \{1^2, 2^3\}$. Set $B = \{a, b\}$ and let Q be the subset of $B^* \times B^*$ consisting of all the following pairs:

(i) $(baaa, abaa)$,

(ii) (bba, bab),

(iii) $(baabaa, ababaa)$,

(iv) $(bababa, babaab)$.

Then (B, Q) *is a finite complete rewriting system defining the monoid* F_1 *where* a *and* b *correspond to the generators* 1^2 *and* 2^3*, respectively.*

2

Preliminaries

2.1 Introduction to semigroups and notation

There are many good publications which provide an introduction to semigroup theory and string rewriting systems. The following are recommended and have been used to source many of the details in this section: [25], [10], [11], [23], [3], [24, §12].

2.2 Semigroups

A *semigroup* S is a set of elements closed under a binary operation which has the additional property of associativity. That is $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in S$. Unlike a group, there is no requirement for an identity element or for each element to have an inverse. If an identity element exists then the semigroup is a *monoid*. A *subsemigroup* T of S is defined as a subset of S which is closed under the binary operation for S.

If a semigroup S contains an element 1 with the property that, for all x in S, $x1 = 1x = x$, we say that 1 is an *identity element* (or just an *identity*) of S, and that S is a *semigroup with identity* or (more usually) a *monoid*. A semigroup S has at most one identity element. If a semigroup has no identity then we can adjoin an identity to create a monoid such that for a semigroup S then S^1 denotes the monoid where $S^1 = S \cup \{1\}$ and 1 is the identity, and we define

$1x = x1 = x$ for all x in S and $11 = 1$. We can also adjoin a zero to create S^0 where $S^0 = S \cup \{0\}$ and we have $0a = a0 = 00 = 0$ for all $a \in S$.

An element a in a monoid M is *right invertible* and called a *right unit* if there exists an element $s \in M$ such that $as = 1$. An element $b \in M$ is *left invertible* and called a *left unit* if there exists an element $t \in M$ such that $ta = 1$. An element $u \in M$ is *invertible* and called a *unit* if there exists an element $v \in M$ such that $uv = vu = 1$. The set of units of a monoid M is a submonoid and is denoted $U(M)$. An element $e \in S$ is an *idempotent* if $ee = e$.

A *right congruence* on a semigroup S is an equivalence relation ρ that is preserved under multiplication on the right. In other words, for all $a, s, t \in S$ we have

$$(s, t) \in \rho \Rightarrow (sa, ta) \in \rho.$$

A *left congruence* on a semigroup S is an equivalence relation ρ that is preserved under multiplication on the left. Hence for all $a, s, t \in S$ we have $(s, t) \in \rho \Rightarrow (as, at) \in \rho$. A relation that is both a left and right congruence is called a *(two-sided) congruence*. If ρ is a congruence on S then we can use a/ρ to denote the congruence class of a and define a (well-defined) binary operation on the quotient set S/ρ by $(a/\rho)(b/\rho) = (ab/\rho)$.

A subsemigroup T of a semigroup S is called a *right ideal* if it satisfies $TS \subseteq T$. Dually, T is called a *left ideal* if $ST \subseteq T$ and a *(two-sided) ideal* if it is both a left and right ideal. An ideal I of S is called a *proper* ideal if $I \neq S$. Note that any left, right or (two-sided) ideal is automatically a subsemigroup. A (left, right or two-sided) ideal I of a semigroup S is said to be *minimal* if it contains no other (left, right or two-sided) ideals of S.

An element a of a semigroup S is called *regular* if there exists $x \in S$ such that $axa = a$. A semigroup S is called a *regular semigroup* if all its elements are regular. Note that all groups are regular semigroups.

A map $\phi : S \to T$ where S and T are semigroups is called a *homomorphism* if for all $x, y \in S$ we have $(xy)\phi = (x\phi)(y\phi)$.

- An *epimorphism* is a surjective homomorphism (also sometimes called an onto homomorphism).
- A *monomorphism* is an injective homomorphism (also sometimes called a 1:1 homomorphism).
- An *isomorphism* is a bijective homomorphism (i.e. both onto and 1:1).
- Two semigroups S and T are said to be isomorphic if there exists an isomorphism $f : S \to T$; this is denoted $S \cong T$.

Definition 2.2.1. [23, Chapter 1, Section 2]

The operation $S \times S \to S$ on a semigroup is usually written like multiplication i.e. $(x, y) \mapsto xy$. The *opposite operation op* on S is defined by $x\ op\ y = yx$ and the resulting semigroup is the *opposite* or *dual* semigroup S^{op}.

Definition 2.2.2. Let $(S, \cdot)$ and $(T, \circ)$ be semigroups, then they are *anti-isomorphic* if there exists some $\phi : S \to T$ such that $(x \cdot y)\phi = (y\phi) \circ (x\phi)$. In other words, there exists an isomorphism from S to the opposite of T.

2.3 Presentations

First let us consider the generators for a semigroup. Let S be a semigroup and let A be a non-empty set which generates S. This means that A is a subset of S such that every element of S can be written as a product of elements from A. If A can be chosen to be finite then S is finitely generated. Then A^+ is the *free semigroup* on A under the operation of concatenation and it consists of all the possible non-empty words which are combinations of the letters from A. An element w of A^+ is called a *word* over A. The number of letters in a word is called the *length of the word* and written as $|w|$. For example, let $w \in A^+$ with $w = a_1a_2 \ldots a_n$ and $a_1, a_2, \ldots, a_n \in A$, then $|w| = n$. We define A^* as the *free monoid* on A with $A^* = A^+ \cup \{\epsilon\}$, where ϵ is the empty word.

Let A be a non-empty set which we call an *alphabet.* A *semigroup presentation* is an ordered pair $\langle A \mid R\rangle$ where R is a set of *relations* such that $R \subseteq A^+ \times A^+$. Relations in the set R can be written as $(u, v) \in R$ with $u, v \in A^+$. The expression (u, v) which we often alternatively write as $(u = v)$ indicates that the word u can be replaced by the word v wherever it appears within a word (this will be explained more formally below).

It is implicit that the letters in an alphabet (as described above) are also elements in the semigroup which they generate. In the case of the free semigroup A^+ our presentation has no defining relations and instead of writing $\langle A \mid \rangle$ or $\langle A\rangle$ we write A^+. Similarly for the free monoid A^*.

Let $A = \{a_1, \ldots, a_m\}$ and $R = \{u_1 = v_1, \ldots, u_n = v_n\}$ then we write $\langle a_1, \ldots, a_m \mid u_1 = v_1, \ldots, u_n = v_n\rangle$ for $\langle A \mid R\rangle$. At this point the presentation is merely a set of symbols. The *semigroup defined by the presentation* $\langle A \mid R\rangle$ is A^+/η where η is the smallest congruence on A^+ containing R. Now we can say that a semigroup S is defined by the presentation $\langle A \mid R\rangle$ providing S is isomorphic to A^+/η, then we write $S \cong A^+/\eta$. We say that the relations R invoke a congruence on the words and generally in this document we will use η to denote the smallest congruence on S containing R. Given any two words $u, v \in A^+$ we write $u \equiv v$ if they are equal as words, and write $u = v$ if they represent the same element of S i.e. if $u/\eta = v/\eta$.

Let $\langle A \mid R\rangle$ be a semigroup presentation, and let $w_1, w_2 \in A^+$ be any words. We say that w_2 is obtained from w_1 by an application of a relation from R if $w_1 \equiv \alpha u\beta$ and $w_2 \equiv \alpha v\beta$, where $\alpha, \beta \in A^*$ and $(u, v) \in R$ or $(v, u) \in R$. We say that the relation $w_1 = w_2$ is a *consequence* of R if w_1 and w_2 are identical words or if there is a sequence $w_1 = \gamma_1, \gamma_2, \ldots, \gamma_m = w_2$ in which each γ_{k+1} is obtained from γ_k by an application of a relation from R. Two words w_1 and w_2 in A^+ are η related if and only if one can transform w_1 into w_2 by a finite number of applications of the defining relations R.

Under this notation the elements of S are in one to one correspondence with the congruence classes of A^+ under η. This means that each word in A^+ *represents* an element in S and each element in S can be represented by one or more words in A^+.

Given any $w \in A^+$ we say that $w/\eta \in S$ is the element of S represented by the word w. For any subset $Y \subseteq S$ we set

$$\mathcal{L}(A, Y) = \{w \in A^+ : w/\eta \in Y\}$$

to be the set of all words in A^+ that represent an element of Y.

Example 2.3.1. Let $S \cong \langle A \mid R \rangle$ where $A = \{a\}$ and $R = \{(aaa = a)\}$ then there is an infinite set of words which represent elements in S but only a finite set of elements. In fact there are only two elements, one which is represented by the set of words $\{a^m : m \geq 1$ and an odd number $\}$ and the other which is represented by the set of words $\{a^n : n \geq 0$ and an even number $\}$.

Theorem 2.3.2. *Let S be a semigroup, let A be a non-empty set and let $f : A \to S$ be any mapping. Then there exists a unique homomorphism $\phi : A^+ \to S$ such that $af = a\phi$ for all $a \in A$. If imf is a generating set of S, then $im\phi$ is S. Therefore, every finitely generated semigroup is a homomorphic image of a finitely generated free semigroup. In fact, every semigroup is isomorphic to a quotient of a free semigroup.*

Theorem 2.3.2 will be applied in the rewriting method described in Section 2.9 and also in Chapter 3.

We say that $\langle A \mid R \rangle$ is a *finite presentation* if both A and R are finite. If a semigroup S is isomorphic to the semigroup defined by some finite presentation then S is said to be *finitely presented.*

Two presentations $\langle A \mid R \rangle$ and $\langle A \mid Q \rangle$, on the same alphabet, are called *equivalent* if they generate the same congruence. Let ρ be the smallest congruence on A^+ containing Q. Two presentations $\langle A \mid R \rangle$ and $\langle B \mid Q \rangle$ are said to be isomorphic to each other if the semigroups A^+/η and B^+/ρ are isomorphic.

2.4 Rewriting systems

Let S be the semigroup defined by the semigroup presentation $\langle A \mid R\rangle$ then there exists a semigroup *rewriting system* which defines S which we denote by (A, R). Here a semigroup rewriting system is a pair (A, R), where A is a non-empty set called an *alphabet*, and $R \subseteq A^+ \times A^+$, called the set of *rewriting rules*. A rewriting system may also be referred to as a *string rewriting system*.

In the terminology of rewriting systems, if (A, R) is a rewriting system then we say that R is a rewriting system over A. Then R comprises a set of rewriting rules where the elements of R are represented as (u, v) or $u \to_R v$ with $(u, v) \in A^+ \times A^+$. Let $(u, v) \in R$ be a rewriting rule in R and w_1, w_2 be words in A^+, with $\alpha, \beta \in A^*$ such that $w_1 \equiv \alpha u \beta$ and $w_2 \equiv \alpha v \beta$. Then we say that w_1 is *rewritten as* (or *reduced to*) w_2 by a one-step *reduction* induced by R which is denoted $w_1 \to_R w_2$. When the set of rewrite rules R is obvious by the context then we simplify notation and write $w_1 \to w_2$.

We use $\xrightarrow{*}_R$ to denote the reflexive transitive closure of $\to_R$. In other words, let $u \xrightarrow{*}_R v$, then we have a finite sequence of words $w_0, w_1, \ldots, w_n$ such that $u \equiv w_0 \to_R w_1 \to_R \ldots \to_R w_n \equiv v$. Note $u \xrightarrow{*}_R v$ implies a finite sequence of zero, one or many rewrite steps. Whereas $u \xrightarrow{+}_R v$ has a similar meaning but implies a finite sequence of at least one rewrite step.

Similarly $\overset{*}{\leftrightarrow}_R$ denotes the reflexive symmetric transitive closure of $\to_R$. In this case , if $u \xleftrightarrow{*}_R v$, then we have a finite sequence of words $w_0, w_1, \ldots, w_n$ such that $u \equiv w_0 \leftrightarrow_R w_1 \leftrightarrow_R \ldots \leftrightarrow_R w_n \equiv v$, where $u \leftrightarrow_R v$ if and only if $u \to_R v$ or $v \to_R u$. Note $u \xleftrightarrow{*}_R v$ implies a finite sequence of zero, one or many rewrite steps. If we now consider the semigroup S, it can be shown that $\overset{*}{\leftrightarrow}_R$ is an equivalence relation on A^+ and it partitions A^+ into congruence classes, in fact the semigroup S defined by the presentation $\langle A \mid R\rangle$ is the quotient semigroup

$A^+/\overset{*}{\leftrightarrow}_R$.

Sometimes it is convenient to use the notation $r = (r_{+1}, r_{-1})$ for a rewrite rule $r \in R$. Then we say that $r_{+1} \to_R r_{-1}$ is a positive elementary transformation. This process of replacing a subword r_{+1} by a word r_{-1} is called a single step reduction. We assume throughout the thesis that if $(r_{+1}, r_{-1}) \in R$ then $(r_{-1}, r_{+1}) \notin R$. If $x, y \in A^+$ and $x \overset{*}{\to} y$, then x is an *ancestor* of y and y is a *descendant* of x.

So if we have a semigroup S defined by the semigroup rewriting system (A, R) then $S \cong A^+/\overset{*}{\leftrightarrow}_R$ and the elements of S are the congruence classes induced by $\overset{*}{\leftrightarrow}_R$. Using the words in A^+ we can represent the elements of S by saying that for $w \in A^+$ then $[w]_R \in S$ where $[w]_R$ denotes the congruence class of w modulo $\overset{*}{\leftrightarrow}_R$.

In string rewriting systems the symbol $\overset{*}{\leftrightarrow}_R$ is also called the Thue congruence on A^+. As for semigroup presentations, the symbol η will be used to denote the smallest congruence on A^+ which contains the relations R. Thus, the following expressions are equivalent and all show that the two words u and v represent the same element of S : $u = v$, $u =_R v$, $[u]_R = [v]_R$, $u \overset{*}{\longleftrightarrow}_R v$, $u/\eta = v/\eta$, $u/\overset{*}{\longleftrightarrow}_R = v/\overset{*}{\longleftrightarrow}_R$. Note that none of these expressions imply $u \equiv v$.

It is worth noting that there is very little difference between the definitions of a semigroup presentation and that for a rewriting system. When considering rewriting systems the emphasis is often on the words and the orientation of rewrite rules is important i.e. $u \to_R v$ where $u, v \in A^+$. Whereas for semigroup presentations the orientation of the relations is not important and we would have $u = v$ instead of $u \leftrightarrow_R v$.

Two string rewriting systems (A, R) and (A, Q) on the same alphabet are called *equivalent* if they generate the same Thue congruence, that is $\overset{*}{\leftrightarrow}_R = \overset{*}{\leftrightarrow}_Q$. Two string rewriting systems (A, R) and (B, Q) are said to be *isomorphic* if the semigroups $A^+/\overset{*}{\leftrightarrow}_R$ and $B^+/\overset{*}{\leftrightarrow}_Q$ are isomorphic to each other.

Let $\mathrm{IRR}(R)$ be the set of all words in A^+ that cannot be reduced by a rule in

R. That is if $w \in \mathrm{IRR}(R)$ then there does not exist a u where $w = A^*uA^*$ and $u \in \mathrm{Left}(R)$ where $\mathrm{Left}(R) = \{u \in A^+ : u \to v \in R\}$. Then if $w \in \mathrm{IRR}(R)$ we say that w is an *irreducible word* and it is not hard to see that we have $\mathrm{IRR}(R) = A^+ \backslash A^* \, \mathrm{Left}(R) A^*$.

2.5 Semigroup vs monoid for presentations and rewriting systems

So far in this chapter we have been concerned with semigroup presentations when referring to the presentation $\langle A \mid R \rangle$. We can make a similar definition with respect to a monoid. A monoid presentation is a pair $\langle A \mid R \rangle$ where A is a non-empty alphabet and R is a subset of $A^* \times A^*$. The monoid presentation $\langle A \mid R \rangle$ defines the monoid A^*/η where η is the smallest congruence on A^* containing R. Just as for semigroup presentations, for any two words u and v we have $u/\eta = v/\eta$ if and only if we can transform u into v by a finite number of applications of the defining relations. Where the context is clear the phrase *monoid presentation* and *semigroup presentation* may be simplified to *presentation*.

As we can see, the two types of presentation are closely related. For example, every semigroup presentation is a monoid presentation as well. If S is the semigroup defined by $\langle A \mid R \rangle$, then the monoid defined by $\langle A \mid R \rangle$ is S with an identity adjoined to it. If $\langle A \mid R \rangle$ is the presentation for the semigroup S and if S possesses an identity $\epsilon \in A^+$, then $\langle A \mid R, \epsilon = 1 \rangle$ is a monoid presentation for S. If M is the monoid defined by the monoid presentation $\langle B \mid G \rangle$, then M can be defined as a semigroup by $\left\langle B, \epsilon \mid G', \epsilon^2 = \epsilon, \epsilon b = b\epsilon = b (b \in B) \right\rangle$, where G' is obtained from G by replacing every relation of the form $w = 1$ by the relation $w = \epsilon$. The following lemma will be used in later chapters.

Lemma 2.5.1. *Let M be the monoid defined by the monoid presentation $\langle A \mid R \rangle$. Then there exists a semigroup presentation which defines M. Moreover, if the monoid presentation is finite, then the semigroup presentation is finite.*

Proof. Let $A' = A \cup \{1\}$ and $R' = R \cup \{(1a, a), (a1, a), (11, 1) : a \in A\}$ where 1 is the identity element. Then it can easily be seen that $\langle A' \mid R' \rangle$ is the semigroup presentation for the monoid M. The rest follows from the definitions of A' and R'. $\square$

Previously in this section we have defined a semigroup rewriting system but we can make a similar definition with respect to a monoid rewriting system. A monoid rewriting system is a pair (A, R), where A is a non-empty alphabet and $R \subseteq A^* \times A^*$ is a set of rewriting rules. The monoid rewriting system (A, R) defines the monoid M where $M \cong A^*/ \xleftrightarrow{*}_R$. Where the context is clear the phrase *monoid rewriting system* and *semigroup rewriting system* may be simplified to *rewriting system.*

Example 2.5.2. The bicyclic monoid is expressed as $\mathbf{B} \cong \langle b, c \mid bc = 1 \rangle$. Let η be the smallest congruence on $\{b, c\}^*$ containing $(bc, 1)$, so $\mathbf{B} \cong \{b, c\}^*/\eta$. It may be shown that each η-class contains exactly one word of the form $c^m b^n$ where $m, n \geq 0$, and hence the elements of $\mathbf{B}$ are represented by this set of words. In terms of words, we could have an infinite number of words which represent the same element. For example $cb = bccb = bcbccb = bbccbccb$ and so on.

Research in this thesis is with respect to both semigroups and monoids. Chapters 3,4 and 5 contain results mostly for semigroups whereas Chapter 6 concentrates on monoids. The context will be made clear at the start of each chapter and within definitions; any differences will be highlighted where relevant.

2.6 Finite complete rewriting systems (FCRS)

Since most of the literature which we make reference to in this subsection is with respect to monoid rewriting systems we will use the convention that the terms *rewriting system* and *string rewriting system* will be referring to *monoid rewriting system* and *monoid string rewriting system* respectively. Note that there are analogous results with respect to semigroup rewriting systems and semigroup string rewriting systems.

A rewriting system (A, R) is called *noetherian* if there is no infinite sequence $w_1, w_2 \ldots$ of words in A^+ such that $w_1 \to w_2 \to \ldots$ for all $w_i \to w_{i+1}, i > 0$. This means that in a noetherian rewriting system any sequence of reducing a word by rewrite rules will eventually terminate at an irreducible word.

Definition 2.6.1. [3, Theorem 2.2.1] Let A be a finite alphabet and $>$ a binary relation on A^*.

(a) The relation $>$ is a *strict partial ordering* if it is irreflexive, anti-symmetric and transitive.

(b) $>$ is a *linear ordering* if it is a strict partial ordering and if, for all $x, y \in A^*$, either $x > y$, or $x = y$, or $y > x$ holds.

(c) The relation $>$ is *admissible* if, for all $u, v, x, y \in A^*$, $u > v$ implies $xuy > xvy$.

Definition 2.6.2. [3, Theorem 2.2.2] Let $A = \{a_1, \ldots, a_n\}$. The following gives some examples of admissible partial orderings on A^*.

(a) Define $x > y$ as follows: $x > y$ if $|x| > |y|$. Then $>$ is the *length ordering* on A^*.

(b) Let $w : A \to \mathbb{N}$ be a mapping that associates a positive integer (a *weight*) with each letter. Define the *weight ordering* $>_w$ induced by w as follows: for x and y in A, $x >_w y$ if $(x)w > (y)w$. Here w is extended to a mapping from A^* into $\mathbb{N}$ by taking $(\epsilon)w := 0$ and defining by induction $(xa)w := (x)w + (a)w$ for all $x \in A^*$, $a \in A$. Note that the length ordering is a weight ordering.

(c) The *lexicographical ordering* $>_{lex}$ on A^* is defined as follows: $x >_{lex} y$ if there is a non-empty string z such that $x = yz$, or $x = ua_iv$ and $y = ua_jz$ for some $u, v, z \in A^*$, and $i, j \in \{1, \ldots, n\}$ satisfying $i > j$.

(d) The *length-lexicographical ordering* $>_{ll}$ is a combination of the length ordering and the lexicographical ordering. Also often referred to as *shortlex ordering* $>_{sh}$:
$x >_{sh} y$ if $|x| > |y|$ or ($|x| = |y|$ and $x >_{lex} y$).

Definition 2.6.3. [3, Theorem 2.2.3] Let $>$ be a strict partial ordering on A^*. It is called *well-founded* if there is no infinite chain of the form $x_0 > x_1 > x_2 > \ldots$. If $>$ is linear and well-founded, then it is called *well-ordering.*

Any weight ordering is well-founded. Since there are only finitely many strings of any given length, the shortlex ordering (also known as length-lexicographical ordering) is a well-ordering. However, if A contains more than one letter, then the lexicographical ordering $>_{lex}$ is not well-founded.

Example 2.6.4. Let $Y = \{x_1 < x_2 < x_3\}$. Then with $<_{sh}$ denoting the shortlex order on words we have $x_1x_2x_3 <_{sh} x_1x_3x_2$ and $x_1x_2 <_{sh} x_1x_1x_2$. Note that in contrast we have $x_1x_1x_2 <_{lex} x_1x_2$. Indeed, lexicographical ordering is not well-founded since

$$x_1x_2 >_{lex} x_1x_1x_2 >_{lex} x_1x_1x_1x_2 >_{lex} \cdots$$

Example 2.6.5. [2, Chapter 2] Examples of well-founded orderings.
(i) Natural numbers with greater than, expressed as $(\mathbb{N}, >)$.

(ii) Let $(M_1, >_1)$ and $(M_2, >_2)$ be well-founded orderings. Then let their lexicographic combination $> = (>_1, >_2)_{lex}$ on $M_1 \times M_2$ be defined as
$(a_1, a_2) > (b_1, b_2) \;\Leftrightarrow\; a_1 >_1 b_1$, or else $a_1 = b_1$ and $a_2 >_2 b_2$
(analogously for more than two orderings). This again yields a well-founded ordering.

This leads us to the following useful lemma.

Lemma 2.6.6. *[2, Chapter 2]* $(M_i, >_i)$ *is well-founded for* $i = 1, 2$ *if and only if* $(M_1 \times M_2, >)$ *with* $> = (>_1, >_2)_{lex}$ *is well-founded.*

Theorem 2.6.7. *[3, Theorem 2.2.4] Let R be a string rewriting system on A. Then the following two statements are equivalent:*

(a) the reduction relation $\to_R$ is noetherian;

(b) there exists an admissible well-founded strict partial ordering $>$ on A^* such that $u > v$ holds for each rule $(u, v) \in R$.

A rewriting system is called *confluent* if, for any words $w, w_1, w_2 \in A^+$ with $w \xrightarrow{*} w_1$ and $w \xrightarrow{*} w_2$, there exists a word $w_0 \in A^+$ such that $w_1 \xrightarrow{*} w_0$ and $w_2 \xrightarrow{*} w_0$. The system is called *locally confluent* if, for any words $w, w_1, w_2 \in A^+$ with $w \to w_1$ and $w \to w_2$, there exists a word $w_0 \in A^+$ such that $w_1 \xrightarrow{*} w_0$ and $w_2 \xrightarrow{*} w_0$.

The following theorem and lemma will be referenced in later chapters.

Theorem 2.6.8. *[42, Newman's Lemma] [3, Theorem 1.1.13] Let $S = (A, R)$ be a string rewriting system which is noetherian. Then S is confluent if and only if S is locally confluent.*

Lemma 2.6.9. *[24, Lemma 12.17] Let (A, R) be a string rewriting system. The system R is locally confluent if and only if the following conditions are satisfied for all pairs of rules $(u_1, t_1), (u_2, t_2) \in R$.*

(i) If $u_1 \equiv rs$ and $u_2 \equiv st$ with $r, s, t \in A^*$ and $s \not\equiv \epsilon$ (the empty word), then there exists $w \in A^*$ with $t_1 t \to^* w$ and $rt_2 \to^* w$.

(ii) If $u_1 \equiv rst$ and $u_2 \equiv s$ with $r, s, t \in A^*$ and $s \not\equiv \epsilon$ (the empty word), then there exists $w \in A^*$ with $t_1 \to^* w$ and $rt_2t \to^* w$.

A pair of rules satisfying either of the two conditions in the above lemma is called a *critical pair*. If R is noetherian and one of these conditions fail (so R is not locally confluent), we end up with two distinct strings, say w_1 and w_2, that are irreducible and equivalent under $\xleftrightarrow{*}$. If all critical pairs in R satisfy conditions (i) and (ii) above then we say that the critical pairs *resolve*.

Next we define what is required for a rewriting system given by the presentation

$\langle A \mid R\rangle$ to be complete. We say that R is a *complete rewriting system* over A if R is both noetherian and confluent. From the results above we deduce:

Lemma 2.6.10. *[24, Lemma 12.15] Suppose that R is noetherian and locally confluent, then R is noetherian and confluent: that is, R is complete.*

Let $\langle A \mid R\rangle$ be a presentation which defines the semigroup S with A and R finite and R complete, then (A, R) is a finite complete semigroup rewriting system. We can then say that (A, R) is a *finite complete rewriting system* which represents the semigroup S and that $\langle A \mid R\rangle$ is a presentation which has the property of being a finite complete rewriting system. We use FCRS as shorthand for a presentation which has the property of being a finite complete rewriting system. Also, we may say that $\langle A \mid R\rangle$ is a complete presentation if (A, R) is complete or that $\langle A \mid R\rangle$ is a finite complete presentation if (A, R) is finite and complete. Most commonly in this thesis we will use the phrase $\langle A \mid R\rangle$ *is a presentation which is a finite complete rewriting system* to mean that the rewriting system (A, R) is finite and complete.

The property of being a finite complete rewriting system is specific to the presentation given. If two presentations describe the same semigroup and one is a finite complete rewriting system, then it does not imply that the second presentation is also a finite complete rewriting system, see below for an example. This is in contrast to some properties and it is this fact which makes it difficult to prove inheritance when considering a substructure of the semigroup.

Examples:

- Let $S \cong \langle A \mid R \rangle = \big\langle a \mid (a^3 \to a) \big\rangle$. Then R is a length reducing system which is a FCRS.
- Let $S \cong \langle A \mid R \rangle = \langle a, b \mid (ba \to ab) \rangle$ is the free commutative monoid of rank 2. This is also a FCRS.
- An example of a semigroup which is finite and complete with respect to one rewriting system but not for another can be found in the paper [27]. In fact, in [27] they prove even more than this, as follows:

  Let $S \cong \langle a, b \mid aba \to bab \rangle$ which the paper proves is not a FCRS. To see this we consider the word *ababa* which can be reduced in two different ways. If we apply the rewrite rule to the first *aba* in the word we get *babba* and if we choose the second occurrence we get *abbab*. Now both these words cannot be reduced further and all three words are equal under the presentation. This means that we do not have a confluent system and thus not a FCRS. In [27] it is shown that S does not have any FCRS with respect to the generating set $\{a, b\}$. However, by adding a further symbol and extending the relations, then the same semigroup can be presented as

  $S \cong \langle a, b, c \mid ab \to c, ca \to bc, bcb \to cc, ccb \to acc \rangle$ which is proved to be a FCRS.

Definition 2.6.11. [3, Definition 1.1.5] Let (A, R) be a string rewriting system. For $x, y \in A^+$, if $x \overset{*}{\leftrightarrow} y$ and y is irreducible, then y is a *normal form* for x.

Theorem 2.6.12. *[14, Proposition 4.1] Let S be a semigroup generated by a set A, let $R \subseteq A^+ \times A^+$ and let $L \subseteq A^+$ be a set of unique normal forms for S. If the following conditions hold:*

(i) S satisfies all the relations from R ; and

(ii) any word $w \in A^+$ can be transformed to the corresponding unique normal form in L by applying relations from R ;

then $\langle A \mid R \rangle$ is a presentation for S.

Theorem 2.6.13. *[3, Theorem 1.1.12] Let (A, R) be a rewriting system for the semigroup S. If (A, R) is complete, then for every $w \in A^*$, $[w]_R$ has a unique normal form.*

Remark: Lemma 2.6.13 relates to the word problem. The word problem for a presentation $\langle A \mid R \rangle$, asks whether there is an algorithm that given any two words u, v over the alphabet A can decide whether or not $u = v$ in the monoid defined by this presentation. That is, whether u can be transformed into v by applying the defining relations. In general there are finitely presented semigroups for which there is no algorithm to solve the word problem. However, if $\langle A \mid R \rangle$ is a presentation which is a finite complete rewriting system then there is an algorithm to solve the word problem. Indeed, using the fact that the rewriting system is complete it follows that applying the rewrite rules to u and v we can rewrite each of them to their unique irreducible forms, call these $\bar{u}$ and $\bar{v}$, then $u = v$ in the monoid if and only if $\bar{u} \equiv \bar{v}$ i.e. the irreducible words are identically equal as words. This lemma implies that we have a rewriting system in which the word problem can be solved.

Lemma 2.6.14. *Let (Z, R) be a finite noetherian rewriting system. Then the following are equivalent:*
(i) (Z, R) is confluent;
(ii) (Z, R) is locally-confluent;
(iii) every $\overset{*}{\leftrightarrow}_R$-class contains exactly one irreducible word.

Proof. The proof is given for each part:

(ii) $\Rightarrow$ (i) If (Z, R) is noetherian and locally confluent then it is also confluent by Theorem 2.6.8.

(i) $\Rightarrow$ (ii) If (Z, R) is noetherian and confluent then it is clearly locally-confluent since confluence clearly implies local confluence.

(ii) $\Rightarrow$ (iii) If (Z, R) is noetherian and locally-confluent then it is complete by Lemma 2.6.10. If (Z, R) is complete then by Theorem 2.6.13 it follows that part (iii) is true.

(iii) $\Rightarrow$ (ii) Suppose that (iii) holds and let $x, u, v \in Z^+$ such that $x \to_R u$ and $x \to_R v$. As (Z, R) is noetherian then $u \xrightarrow{*}_R u'$ and $v \xrightarrow{*}_R v'$ for some $u', v' \in \mathrm{IRR}(R)$. Then (iii) implies $u' \equiv v'$ and therefore (Z, R) is locally-confluent. □

Definition 2.6.15. A rewriting system is said to be *normalised* if for every $u \to v \in R$ we have $v \in \mathrm{IRR}(R)$ and there does not exist a $u' \in A^+$ with $u \to u'$ by any other rule in $R \setminus (u \to v)$.

Note that complete string rewriting systems that are also normalised are called *canonical*.

Lemma 2.6.16. *[3, Algorithm 2.2.12] Let (A, R') be a finite complete rewriting system. Then there is a finite complete rewriting system (A, R) which is equivalent to (A, R') such that R is normalised.*

It is a key fact that when we apply Lemma 2.6.16, this normalising procedure does not change the generating set.

Lemma 2.6.17. *Let (A, R) be a finite complete rewriting system. Let $w \in \mathrm{IRR}(R)$, $u \in A^+$, $w_1, w_2 \in A^*$, then*

(i) Any subword u of w is also irreducible i.e. whenever $w \equiv w_1 u w_2$ then $u \in \mathrm{IRR}(R)$.

(ii) For any proper subword u of w we cannot have $u \xleftrightarrow{*}_R w$.

Proof. (i) Suppose for a contradiction $w \equiv w_1 u w_2$, $w \in \mathrm{IRR}(R)$ and $u \notin \mathrm{IRR}(R)$. Then we have $u \to_R v \in R$ and hence $w_1 u w_2 \to_R w_1 v w_2$. This implies $w \equiv w_1 u w_2 \notin \mathrm{IRR}(R)$, a contradiction.

(ii) Let $w \equiv w_1 u w_2$ with $u \not\equiv w$. Suppose for a contradiction we have $u \xleftrightarrow{*}_R w$ which means that u and w represent the same element i.e. $u/\eta = w/\eta$. As

$w \in \mathrm{IRR}(R)$ and $u \not\equiv w$ and we have a finite complete rewriting system, then there exists a sequence of rewrite rules such that $u \xrightarrow{+}_R w$. But this means that $u \notin \mathrm{IRR}(R)$ which contradicts (i). $\square$

We now list some results which show that many of the properties discussed above are shared between S and T if T is a subsemigroup of S that differs from S by only finitely many elements. There are several expressions which all convey this situation, namely

(i) T has finite Rees index in S;

(ii) S is a small extension of T;

(iii) T is a large subsemigroup of S;

(iv) $S \setminus T$ is finite.

We choose to use the last expression in the following theorems and corollary.

Theorem 2.6.18. *[48, Theorem 1.3] Let S be a semigroup and T be a subsemigroup of S such that $S \setminus T$ is finite. Then S is finitely presented if and only if T is finitely presented.*

Theorem 2.6.19. *[53, Theorem 1] Let S be a monoid and T be a submonoid of S such that $S \setminus T$ is finite. If T can be presented by a finite complete rewriting system, then so can S.*

Theorem 2.6.20. *[54, Theorem 1.1] Let S be a semigroup and T be a subsemigroup of S such that $S \setminus T$ is finite. If S has a finite complete rewriting system, then so does T.*

Corollary 2.6.21. *[54, Corollary 1.2] Let S be a semigroup and T be a subsemigroup of S such that $S \setminus T$ is finite. Then S has a finite complete rewriting system if and only if T does.*

Although we may be considering a monoid and a submonoid in terms of the property FCRS, we can translate any result to that for semigroups and subsemigroups by appealing to a result in the paper [17] which follows.

Theorem 2.6.22. *[17, Theorem 1.1] Let S be a monoid. Then S is defined by a finite complete semigroup presentation if and only if it is defined by a finite complete monoid presentation.*

2.7 Green's relations

A general method for analysing the structure of a semigroup is to look at its Green's relations (introduced in 1951 by J.A. Green as described in [22]). Let S be a semigroup, then two elements $s, t \in S$ are said to be $\mathcal{L}$-related (denoted $a\mathcal{L}b$) if they generate the same left ideal i.e. $S^1 s = S^1 t$. In other words, there exist $a, b \in S^1$ such that $as = t$ and $bt = s$. The relation $\mathcal{L}$ is an equivalence relation on S and a right congruence. Similarly, two elements are said to be $\mathcal{R}$-related (denoted $s\mathcal{R}t$) if they generate the same right ideal i.e. $sS^1 = tS^1$. In other words, there exist $a, b \in S^1$ such that $sa = t$ and $tb = s$. The relation $\mathcal{R}$ is an equivalence relation on S and a left congruence.

When two elements are in the same $\mathcal{L}$-class and the same $\mathcal{R}$-class, they are said to be $\mathcal{H}$-related. The smallest equivalence containing both $\mathcal{L}$ and $\mathcal{R}$ is called the $\mathcal{D}$-relation. It can be shown that $\mathcal{L}$ and $\mathcal{R}$ commute and we have $\mathcal{D} = \mathcal{L} \circ \mathcal{R} = \mathcal{R} \circ \mathcal{L}$. Two elements are $\mathcal{J}$-related if $S^1 a S^1 = S^1 b S^1$. See [25] for further details and proofs.

It is useful to look at the structure of the $\mathcal{D}$-classes in a semigroup. Also, each individual $\mathcal{D}$-class has its own structure. To visualise the structure of a $\mathcal{D}$-class a grid like diagram is usually drawn with each $\mathcal{R}$-class, $\mathcal{L}$-class and $\mathcal{H}$-class represented by a row, column or square intersection respectively. This is usually referred to as an egg box diagram, see Figure 2.7.1 below. It is possible for the egg box diagram to contain a single row or a single column of cells, or even to contain only one cell. Also, it may well be an infinite egg box.

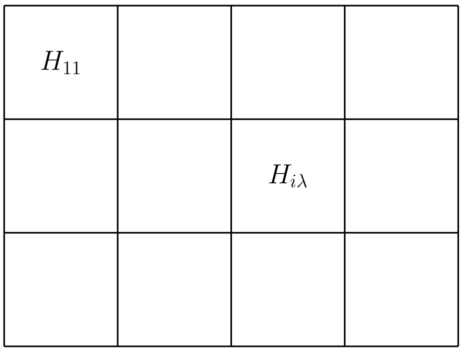


Figure 2.7.1: An egg box diagram for a single $\mathcal{D}$-class

2.8 Tietze transformations

In Section 2.3 we defined the notions of two presentations being equivalent or being isomorphic. Another useful tool is that of changing a presentation by making Tietze transformations. These are simple changes to the generating set or to the set of relations. In this way a new presentation can be created which defines the same semigroup. We say that $u = v$ is a *consequence* of the presentation $\langle A \mid R \rangle$ if $u \overset{*}{\longleftrightarrow}_R v$.

Let $\langle A \mid R \rangle$ be the presentation for a semigroup S. These changes are called *elementary Tietze transformations* and are as follows:

T1 Adding a new relation $u = v$ to $\langle A \mid R \rangle$, provided that $u = v$ is a consequence of $\langle A \mid R \rangle$.

T2 Deleting a relation $u = v \in R$ from $\langle A \mid R \rangle$, provided that $u = v$ is a consequence of $\langle A \mid R \setminus \{(u = v)\} \rangle$.

T3 Adding a new generating symbol b and a new relation $b = w$ for any non-empty word $w \in A^+$.

T4 If $\langle A \mid R \rangle$ possesses a relation of the form $b = w$, where $b \in A$ and $w \in (A \setminus \{b\})^+$, then deleting b from the list of generating symbols, deleting the relation $b = w$, and replacing all remaining appearances of b by w.

Theorem 2.8.1. *[47, Proposition 2.5] Two finite presentations define the same* *semigroup (i.e. they are isomorphic) if and only if one can be obtained from the other by a finite number of applications of elementary Tietze transformations (T1), (T2), (T3), (T4).*

We can also say:

Theorem 2.8.2. *Two finite presentations $\langle A \mid R \rangle$ and $\langle A \mid Q \rangle$ are equivalent if and only if $\langle A \mid R \rangle$ can be obtained from $\langle A \mid Q \rangle$ by a finite number of applications of elementary Tietze transformations (T1) and (T2).*

2.9 Reidemeister-Schreier type rewriting for semigroups

The method outlined in this section was published as part of a research article [5] by C.M. Campbell, E.F. Robertson, N. Ruškuc and R.M. Thomas. The approach is used in chapters of this thesis and this section provides an overview. The following material is taken from [5].

In group theory the theorems of Reidemeister and Schreier give a method for determining a presentation for a subgroup H when given a presentation for the group G. One consequence is that, if the group presentation is finite and H has finite index in G, then H is itself finitely presented. (Note that the *index* of H in G is the number of cosets of H in G.) The research was interested in developing an analogous theory for semigroups with a corresponding situation where there exists a (two-sided) ideal I in a semigroup S such that the Rees quotient S/I is finite (equivalently, such that the complement $S \setminus I$ is finite). A constructive method was determined for finding a presentation for a semigroup T of S.

If S is a semigroup defined by a presentation $\langle A \mid R \rangle$, and we are finding a presentation $\langle B \mid R_1 \rangle$ for a subsemigroup T which is specified by giving a set of words X which generate T, and each letter of B corresponds to one of the words from this generating set X. Then we have a mapping from B to A^+, which naturally extends to a homomorphism from B^+ to A^+. The critical idea here is that of a *rewriting mapping* ϕ which, given a word in A^+ which represents an element of T, rewrites it into a corresponding word in B^+. We will define these ideas more formally next.

A general rewriting theorem

Let S be the semigroup defined by a presentation $\langle A \mid R \rangle$, where A is an alphabet and $R \subseteq A^+ \times A^+$, and let T be the subsemigroup of S generated by a set $X = \{\xi_i \mid i \in I\}$, where ξ_i are words from A^+. In this section we are looking for a general presentation for T in terms of the generators X.

First we introduce a new alphabet $B = \{b_i \mid i \in I\}$ in 1:1 correspondence with the set X, and a homomorphism $\psi : B^+ \to A^+$ induced by

$$(b_i)\psi = \xi_i, \ \ i \in I. \tag{2.9.1}$$

It is obvious that $Im\psi$, when represented as a subset of S, is actually T. Intuitively, ψ interprets each word in B^+ as an element of T; sometimes we will call ψ the *interpretation mapping.* A *rewriting mapping* is a mapping

$$\phi \ : \ \mathcal{L}(A, T) \to B^+ \tag{2.9.2}$$

with the property that

$$\big((w\phi)\psi\big) = w \text{ in } S. \tag{2.9.3}$$

Intuitively, ϕ *rewrites* every element of T as a product of the given generators for T. It is not hard to prove that a rewriting mapping always exists. There will typically be many choices for the rewriting mapping. A key part of using this method effectively is to make a good choice of rewriting mapping. We can now state the result giving a presentation for T.

Theorem 2.9.1. *[5, Theorem 2.1] Let S be the semigroup defined by a* *presentation* $\langle A \mid R \rangle$, *and let* T *be the subsemigroup of* S *generated by* $X = \{\xi_i \mid i \in I\} \subseteq A^+$. *Introduce a new alphabet* $B = \{b_i \mid i \in I\}$, *and let* ψ *and* ϕ *be the interpretation mapping and a rewriting mapping. Then* T *is defined by the generators* B *and the relations*

$$b_i = (\xi_i)\phi, \quad i \in I, \tag{2.9.4}$$

$$(w_1 w_2)\phi = (w_1)\phi(w_2)\phi, \tag{2.9.5}$$

$$(w_3 u w_4)\phi = (w_3 v w_4)\phi, \tag{2.9.6}$$

where $w_1, w_2 \in \mathcal{L}(A, T)$, $u = v$ *is a relation from* R, *and* $w_3, w_4 \in A^*$ *are any words such that* $w_3 u w_4 \in \mathcal{L}(A, T)$.

Remark: The main disadvantage of the presentation (2.9.4), (2.9.5), (2.9.6) is that it is always infinite since typically w_1, w_2, w_3 and w_4 will range over infinitely many different words, giving rise to infinitely many distinct relations in (2.9.5) and (2.9.6), and that it crucially depends on the mapping ϕ which has not been constructively defined. The significance of the theorem is that it gives us a general recipe for finding presentations for semigroups in various special cases. When applying this result we first find a generating set for the subsemigroup in question, then define a specific rewriting mapping, and then seek a smaller set (ideally finite) of relations which imply all the relations (2.9.4), (2.9.5) and (2.9.6).

3

Rees matrix semigroups

3.1 Introduction

In this chapter we will consider finiteness properties of specific Rees matrix semigroups and their substructures. In particular, we will look at completely simple semigroups which are isomorphic to Rees matrix semigroups of the form $S = \mathcal{M}[G; I, \Lambda; P]$ where G is a group and the indexes I and Λ are finite (see Theorem 3.2.6). In this case the substructure we consider is the group $\mathcal{H}$-classes, also known as maximal subgroups. In this chapter we will be interested in problems of the following kind.

Statement [48, Remark and open problem 4.5]:
Let S be a regular semigroup with finitely many $\mathcal{R}$-classes and $\mathcal{L}$-classes. Let P be a property of semigroups. Then S has property P if and only if all maximal subgroups of S have property P.

The case when P is the property of being "finitely generated" and "finitely presented" have been proved in [46, Proposition 4.2] and the property "has finite derivation type" (see Chapter 4 for more details) in [38, Theorem 2]. The case when P is the property "admits a presentation by a finite complete rewriting system", is proved true in one direction in [18, Corollary 4] for the case where S has the property if every maximal subgroup of S has the property.

We are interested in proving the converse for this last property, "admits a presentation by a finite complete rewriting system", in one special case. In this chapter we will consider this question in the case where S is a completely simple semigroup. We will prove a partial converse result for a particular class of Rees matrix semigroups with an additional condition on the generating set called being sparse (see Definition 3.3.2 below). The general case remains open. The main result in this chapter is as follows:

Theorem 3.1.1. *Let $S = \mathcal{M}[G; I, \Lambda; P]$ be a Rees matrix semigroup where G is a group and one of two conditions is true:*
(i) $I = \{1\}$ and $\Lambda = \{1, 2, \ldots, n\}$;
(ii) $I = \{1, \ldots, m\}$ and $\Lambda = \{1\}$.

Then the group G admits a presentation by a finite complete rewriting system if and only if S admits a presentation by a finite complete rewriting system with respect to some sparse generating set.

Before going on to prove this result we first recall some background results and definitions.

3.2 Background and definitions

Definition 3.2.1. [25, Section 3] A semigroup is called *simple* if it has no proper ideals.

Theorem 3.2.2. *[25, Corollary 3.1.2] A semigroup S is simple if and only if $SaS = S$ for all a in S, that is, if and only if for every a, b in S there exist x, y in S such that $xay = b$.*

Corollary 3.2.3. *Let S be a simple semigroup. Then S has a single $\mathcal{J}$-class.*

Proof. The proof is immediate from the definition of a simple semigroup. □

Example 3.2.4. Every group G is a simple semigroup. To see this we take arbitrary $a, b \in G$ and we have $baa^{-1} = b$ as per Theorem 3.2.2.

A (left, right, two-sided) ideal I of a semigroup S is said to be *minimal* if it contains no other (left, right, two-sided) ideals of S.

Definition 3.2.5. A semigroup S is said to be *completely simple* if it is simple and if it possesses minimal left and right ideals.

Theorem 3.2.6. *[25, Theorem 3.3.1] Let G be a group, let I, Λ be non-empty sets and let $P = (p_{\lambda i})$ be a $\Lambda \times I$ matrix with entries in G. Let $S = (I \times G \times \Lambda)$ and define a multiplication on S by*

$$(i, a, \lambda)(j, b, \mu) = (i, ap_{\lambda j}b, \mu).$$

Then S is a completely simple semigroup.
Conversely, every completely simple semigroup is isomorphic to a semigroup constructed in this way. We denote the semigroup $S = (I \times G \times \Lambda)$ with the given multiplication by $\mathcal{M}[G; I, \Lambda; P]$.

A matrix P is *normal* if every entry in the first row and the first column of P is equal to the identity of the group G.

Theorem 3.2.7. *[25, Theorem 3.4.2] If S is a completely simple semigroup then S is isomorphic to a Rees matrix semigroup $\mathcal{M}[G; I, \Lambda; P]$ in which the matrix P is normal.*

Theorem 3.2.8. *[25, Theorem 2.2.5] If H is an $\mathcal{H}$-class in a semigroup S then either $H^2 \cap H = \varnothing$ or $H^2 = H$ and H is a subgroup of S.*

Proposition 3.2.9. *[25, Proposition 2.3.6] If H and K are two group $\mathcal{H}$-classes in the same regular $\mathcal{D}$-class, then H and K are isomorphic.*

The *maximal subgroups* of a semigroup S coincide with the group $\mathcal{H}$-classes.

3.3 Definitions and notation for new research

In this section we establish the specific criteria for the Rees matrix semigroup $S = \mathcal{M}[G; I, \Lambda; P]$ which we will go on to consider. Note that we are working with a semigroup S and therefore with semigroup presentations and semigroup rewriting systems.

Let $S = \mathcal{M}[G; I, \Lambda; P]$ as in Theorem 3.2.6. Furthermore, let us suppose that $I = \{1, \ldots, m\}$ and $\Lambda = \{1, \ldots, n\}$ both finite. By Theorem 3.2.7 we will assume throughout that the matrix P is normal so that every entry in the first row and the first column is equal to the identity of G.

The semigroup S has a single $\mathcal{D}$-class, m $\mathcal{R}$-classes and n $\mathcal{L}$-classes. We will use the following notation to describe this structure. Let R_i, $i \in I$ be the $\mathcal{R}$-class of S and let L_λ, $\lambda \in \Lambda$ be the $\mathcal{L}$-class of S. The $\mathcal{H}$-classes are the intersections of the $\mathcal{L}$- and $\mathcal{R}$-classes. Let $H_{i\lambda}$, $i \in I$, $\lambda \in \Lambda$ denote the $\mathcal{H}$-class which is the intersection of R_i and L_λ. Thus if an element $s \in S$ belongs to R_2 and L_3 then it follows that $s \in H_{23}$. If $s \in S$, then we also use the notation R_s, L_s and H_s to represent the element's $\mathcal{R}$-class, $\mathcal{L}$-class and $\mathcal{H}$-class respectively. For a Rees matrix semigroup where $S = \mathcal{M}[G; I, \Lambda; P]$ with G being a group, each $\mathcal{H}$-class is a group $\mathcal{H}$-class and will contain a single idempotent element. We denote the idempotent in the $\mathcal{H}$-class $H_{i\lambda}$ by $e_{i\lambda}$. Recall that for an idempotent $e_{i\lambda}e_{i\lambda} = e_{i\lambda}$. The $\mathcal{H}$-class $H_{i\lambda}$ will consist of the triples (i, g, λ) where $g \in G$. Every idempotent $e \in S$ is a left identity for its $\mathcal{R}$-class R_e and a right identity for its $\mathcal{L}$-class L_e. So for $e_{i\lambda}$, $s \in H_{ij}$ and $t \in H_{k\lambda}$ then $e_{i\lambda}s = s$ and $te_{i\lambda} = t$.

For further details on Green's relations and their notation see [25].

In order to progress this line of study it was necessary to add further conditions and it was decided to assume that the generating set for the semigroup was of a specific form which we will call a *sparse generating set*. First we identify a set of elements in S which generate the semigroup.

Lemma 3.3.1. *Let $S = \mathcal{M}[G; I, \Lambda; P]$ with G, I, Λ and P as defined above. Let G_A be a finite semigroup generating set for the group G. Let*

$Q = \{(1, g, 1) : g \in G_A\}$ *which is a semigroup generating set for H_{11};*

$Y = \{(1, 1_G, \lambda) : \lambda \in \Lambda \setminus \{1\}\}$ *be the set of idempotents, one from each $H_{1\lambda}$;*

$Z = \{(i, 1_G, 1) : i \in I \setminus \{1\}\}$ *be the set of idempotents, one from each H_{i1}.*

Then $(Q \cup Y \cup Z)$ is a finite generating set for S.

Proof. Let $s = (j, h, \mu)$ be an arbitrary element in S such that $j \in I$, $\mu \in \Lambda$ and $h \in G$. There exists an element $(1, h, 1)$ in $\langle Q \rangle$ since $(1, h, 1)$ is in H_{11} and Q generates H_{11}. Then

$$\begin{aligned} s &= (j, 1_G, 1)(1, h, 1)(1, 1_G, \mu) \\ &= (j, h, 1)(1, 1_G, \mu) \\ &= (j, h, \mu). \end{aligned}$$

Therefore $(Q \cup Y \cup Z)$ is a finite generating set for S. □

We call a generating set of the form given in Lemma 3.3.1 a *sparse generating set*. To be more precise:

Definition 3.3.2. Let $S = \mathcal{M}[G; I, \Lambda; P]$ be a Rees matrix semigroup. We call a subset X of S a sparse generating set for S if $\langle X \rangle = S$ and the following conditions are satisfied:

(i) $\langle X \cap H_{11} \rangle = H_{11}$;

(ii) $X \setminus H_{11}$ is the set of all idempotents in the $\mathcal{R}$-class of H_{11} and the $\mathcal{L}$-class of H_{11}, excluding the idempotent in H_{11}.

It follows from Lemma 3.3.1 that every Rees matrix semigroup admits a sparse generating set. It follows that every Rees matrix semigroup admits a presentation with respect to a sparse generating set. Moreover, if S is finitely generated then S admits a finite sparse generating set, and if S is finitely presented then S admits a finite presentation with respect to a sparse generating set. If such a presentation exists and the associated rewriting system is also complete, then we say S admits a presentation which is a finite complete rewriting system with respect to a sparse generating set. We shall use the following notation and conventions when considering a presentation for S with respect to a sparse generating set (as mentioned above, such a presentation must exist).

Notation 3.3.3. A presentation for S with respect to a sparse generating set will be written as $\langle A \cup B \cup C \mid R \rangle$ where A, B, C are pairwise disjoint sets and where

(i) A is the set of letters which correspond to the elements in the set Q;

(ii) B is the set of letters which correspond to the elements in the set Y;

(iii) C is the set of letters which correspond to the elements in the set Z;

all defined such that there is a bijection between the letters of each alphabet and the elements in the corresponding set. We call $A \cup B \cup C$ the *sparse generating set* for S.

Let $R \subseteq (A \cup B \cup C)^+ \times (A \cup B \cup C)^+$ be a set of relations such that $(A \cup B \cup C)^+ / \overset{*}{\leftrightarrow}_R$ is isomorphic to S. Let η be the smallest congruence on $(A \cup B \cup C)^+$ which contains R. We shall use the notation

(a) $A = \{a_g : g \in G_A\}$ where a_g/η corresponds to $q_g = (1, g, 1) \in Q$.

(b) $B = \{b_\lambda : \lambda \in \Lambda \setminus \{1\}\}$ where b_λ/η corresponds to $e_{1\lambda} = (1, 1_G, \lambda) \in Y$.

(c) $C = \{c_i : i \in I \setminus \{1\}\}$ where c_i/η corresponds to $e_{i1} = (i, 1_G, 1) \in Z$.

3.4 New research regarding the special case of a single $\mathcal{R}$-class

Let us consider the special case for $S = \mathcal{M}[G; I, \Lambda; P]$ where $I = \{1\}$, $\Lambda = \{1, 2, \ldots, n\}$ with G and P as defined in Section 3.3. Our aim is to prove Theorem 3.1.1 part (i), Theorem 3.1.1 part (ii) then follows by a dual argument.

First we look to prove Theorem 3.1.1 (i) in the ($\Leftarrow$) direction.

Therefore our initial assumption is that the semigroup S admits a presentation which is a finite complete rewriting system with respect to a sparse generating set. As $I = \{1\}$ we make some adjustments to the definitions made in Section 3.3 which will simplify the notation and will apply for the remainder of this section.

First we specify our sparse generating set for this special case. Let the sets A and B be as defined in Section 3.3. The set C is empty and can be ignored. We call $A \cup B$ the sparse generating set for the semigroup S as per Definition 3.3.2 and Notation 3.3.3.

Let $R \subseteq (A \cup B)^+ \times (A \cup B)^+$ be a set of relations such that $(A \cup B)^+ / \overset{*}{\leftrightarrow}_R$ is isomorphic to S and the rewriting system $(A \cup B, R)$ is complete. Furthermore, by Lemma 2.6.16 without loss of generality we can, and will, assume that this rewriting system is normalised, in the sense of Definition 2.6.15. Let η be the smallest congruence on $(A \cup B)^+$ which contains R. Simplifying the notation

(a) $A = \{a_g : g \in G_A\}$ where a_g/η corresponds to $q_g = (1, g, 1) \in Q$.

(b) $B = \{b_\lambda : \lambda \in \Lambda \setminus \{1\}\}$ where b_λ/η corresponds to $e_\lambda = (1, 1_G, \lambda) \in Y$.

The following diagram, Figure 3.4.1, illustrates the semigroup S using an egg box diagram. Examples of elements are illustrated within their respective $\mathcal{H}$-classes together with the different notations with respect to the Rees matrix triples and the semigroup presentation $\langle A \cup B \mid R \rangle$.

H_{11}		$H_{1\lambda}$	H_{1n}
•$a_g/\eta \in A/\eta$ $q_g = (1, g, 1)$		•$b_\lambda/\eta \in B/\eta$ $e_\lambda = (1, 1_G, \lambda)$	
•$(a_g a_h)/\eta$ $(1, gh, 1)$		•$q_g e_\lambda = (1, g, \lambda)$	

Figure 3.4.1: The $\mathcal{D}$-class structure for the special case of a single $\mathcal{R}$-class

Now we consider the semigroup presentation $\langle A \cup B \mid R \rangle$ and the associated semigroup rewriting system $(A \cup B, R)$ which define S. We aim to prove that the group G admits a presentation which is a finite complete rewriting system.

The group $\mathcal{H}$-class H_{11} is a subgroup of S which is isomorphic to the group G. This is true by the definition of the Rees matrix semigroup S and of Q which generates H_{11}, see Lemma 3.3.1. In addition, we have defined the set A which corresponds to the elements of Q and therefore A/η is a finite generating set for H_{11}. Let

$$R_H = \{(u, v) \in R : u, v \in A^+\}.$$

We claim that $\langle A \mid R_H \rangle$ is a semigroup presentation which is a finite complete rewriting system for the subsemigroup H_{11} of S and hence also for the group G. First we prove that $\langle A \mid R_H \rangle$ is a finite presentation which defines H_{11} and then that the rewriting system (A, R) is complete.

Next we include some useful lemmas. The first of which characterises the $\mathcal{H}$-class of an element w/η where $w \in (A \cup B)^+$, in terms of the last letter of the word w.

Lemma 3.4.1. *Let* $w \equiv x_1 x_2 \dots x_p \in (A \cup B)^+$ *be an arbitrary word where* $x_d \in (A \cup B)$ *for* $d \in \{1, \dots, p\}$. *Then*

(i) $w/\eta \in H_{11}$ *if and only if* $x_p \in A$, *and*

(ii) $w/\eta \in H_{1\lambda}$ *for some* $\lambda \in \{2, 3, \dots, n\}$ *if and only if* $x_p \equiv b_\lambda \in B$.

Proof. The proof for parts (i) and (ii) follows from the definition of multiplication for elements of Rees matrix semigroups (Theorem 3.2.6) together with the correspondence established in Lemma 3.3.1 and Notation 3.3.3 between the sets A, B and the sets Q, Y. The proof for part (i) follows in detail. A similar proof can be constructed for part (ii).

Part (i) ($\Rightarrow$). Let $w/\eta \in H_{11}$ and we aim to show that $x_p \in A$. For a contradiction, assume $x_p \equiv b_\lambda \in B$ with $\lambda \neq 1$, so we have $w \equiv x_1x_2\ldots b_\lambda$ and let $w \equiv w'b_\lambda$. But then w represents the element $w/\eta = (w'/\eta)(b_\lambda/\eta) = (w'/\eta)e_\lambda$ which must be in the same $\mathcal{H}$-class as $e_\lambda \in H_{1\lambda}$, a contradiction. To see this, recall how multiplication of elements works in S. Let $(w'/\eta) = (1, g, \mu)$, $g \in G$, $\mu \in \Lambda$ and we have $e_\lambda = (1, 1_G, \lambda)$. Then $w/\eta = (w'/\eta)e_\lambda = (1, g, \mu)(1, 1_G, \lambda) = (1, g, \lambda)$ and so e_λ and w/η must be in the same $\mathcal{H}$-class, which is $H_{1\lambda}$, and therefore a contradiction. So we must have $x_p \in A$.

Part (i) ($\Leftarrow$). Let $x_p \equiv a_1 \in A$ and we aim to show $w/\eta \in H_{11}$. Now $w \equiv x_1x_2\ldots a_1$ and let $w \equiv w'a_1$ which gives us $w/\eta = (w'/\eta)(a_1/\eta)$. We again look at multiplication of elements in S. Let $(w'/\eta) = (1, g, \mu)$, $g \in G$, $\mu \in \Lambda$, and let $(a_1/\eta) = (1, h, 1)$, $h \in G_A$ as $a_1 \in A$ and $(a_1/\eta) \in H_{11}$. Then $w/\eta = (w'/\eta)(a_1/\eta) = (1, g, \mu)(1, h, 1) = (1, gh, 1)$ which is an element of H_{11} as required. $\square$

Lemma 3.4.2. *If* $w \in \mathrm{IRR}(R) \cap \mathcal{L}(A \cup B, H_{11})$ *then* $w \in A^+$.

Proof. Suppose, for a proof by contradiction, that $w \notin A^+$. Let $w \equiv w_1bw_2$ with $b \in B$, $w_1 \in (A \cup B)^*$, $w_2 \in A^+$ (by Lemma 3.4.1 (i)), such that b is the rightmost occurrence of a letter from B in w. Then let a be the first letter of the word w_2 so that $w_2 \equiv aw_2'$ where $w_2' \in A^*$. So we have

$$w \equiv w_1baw_2'.$$

Then $(ba)/\eta = (b/\eta)(a/\eta) = e_\lambda(a/\eta)$ where $\lambda \in \{2, \ldots, n\}$ and e_λ is an idempotent. As e_λ is an idempotent it is a left identity for its $\mathcal{R}$-class and therefore $(ba)/\eta = e_\lambda(a/\eta) = a/\eta$. Since w is irreducible (by definition of this lemma) then by Lemma 2.6.17 (i) all of its subwords are irreducible. Since w is irreducible and ba is a subword of w it follows that ba is an irreducible word. So ba is irreducible and $a \overset{*}{\leftrightarrow}_R ba$. But since a is a subword of ba and ba is irreducible this contradicts Lemma 2.6.17 (ii). □

Lemma 3.4.3. *For any rewrite rule $u \to v$ from R, if $u \in A^+$ then $v \in A^+$.*

Proof. If $u \in A^+$, then $u \in H_{11}$ by Lemma 3.4.1 (i). As $u = v$ in the semigroup S it follows that $v \in H_{11}$. Since we have $v \in \mathrm{IRR}(R)$, it follows that $v \in \mathrm{IRR}(R) \cap \mathcal{L}(A \cup B, H_{11})$ and so by Lemma 3.4.2, $v \in A^+$. □

Lemma 3.4.4. *Let $\langle A \mid R_H \rangle$ be the presentation defined above. Then $\langle A \mid R_H \rangle$ is a finite presentation for H_{11}.*

Proof. The proof uses the properties of the rewriting systems (A, R_H) and (A, R). By definition of our sparse generating set (Lemma 3.3.1) we have shown that A/η generates H_{11}. Let $w_1, w^{'}$ be words in $\mathcal{L}(A, H_{11})$ such that $w' \in \mathrm{IRR}(R)$ and $w_1 \overset{*}{\to}_R w'$. We aim to show that $w_1 \overset{*}{\to}_{R_H} w'$. Note that $w' \in A^+$ by Lemma 3.4.2. Let $w_1 \to_R w_2 \to_R \ldots \to_R w_k \equiv w'$ and we first consider $w_1 \to_R w_2$. As $w_1 \in A^+$ then $w_2 \in A^+$ by Lemma 3.4.3. This follows as all the subwords of w_1 are in A^+ and as such can only be rewritten to subwords in w_2 that are also in A^+. Therefore the rewrite rule which takes w_1 to w_2 must be in the subset R_H and so we also have $w_1 \to_{R_H} w_2$. So for each rewrite step $w_i \to_R w_{i+1}$ with $i \in \{1, \ldots, k-1\}$ we will have a corresponding rewrite rule in R_H and therefore the rewrite step $w_i \to_{R_H} w_{i+1}$ exists in the rewriting system (A, R_H). In fact each word w_j with $j \in \{1, \ldots, k\}$ is in A^+. It follows that for any two words $u, v \in \mathcal{L}(A, H_{11})$ we have $u \overset{*}{\leftrightarrow}_R v$ if and only if $u \overset{*}{\leftrightarrow}_{R_H} v$. Together with the fact that A/η generates H_{11} if follows from this that $\langle A \mid R_H \rangle$ is a presentation for H_{11}. Finally, as R is finite, then R_H is finite. Hence $\langle A \mid R_H \rangle$ is a finite semigroup presentation for H_{11}. □

Lemma 3.4.5. *Let (A, R_H) be the rewriting system as defined above. Then it is locally confluent and noetherian and thus complete.*

Proof. First we consider the property of local confluence. Let $w, w_1, w_2 \in A^+$ such that $w \rightarrow_{R_H} w_1$ and $w \rightarrow_{R_H} w_2$ with $w_1 \not\equiv w_2$. Then as $R_H \subseteq R$ and R is locally confluent, there exists $w' \in (A \cup B)^+$ such that $w \rightarrow_{R_H} w_1 \xrightarrow{*}_R w'$ and $w \rightarrow_{R_H} w_2 \xrightarrow{*}_R w'$. Consider the rewriting path $w \rightarrow_{R_H} w_1 \xrightarrow{*}_R w'$ and by Lemma 3.4.3, all the words in the rewrite path must belong to A^+. Also, for each rewrite rule $(u_i, v_i) \in R$ that is applied we have $u_i, v_i \in A^+$ and therefore $(u_i, v_i) \in R_H$. It follows that $w \rightarrow_{R_H} w_1 \xrightarrow{*}_{R_H} w'$ and similarly $w \rightarrow_{R_H} w_2 \xrightarrow{*}_{R_H} w'$. Thus R_H over A^+ is locally confluent.

It is a fact that subsystems of noetherian rewriting systems are noetherian. To see this refer to Theorem 2.6.7. If R is noetherian, Theorem 2.6.7 (b) is true, but then by definition, Theorem 2.6.7 (b) is true for any subset of R and therefore any subset of R is also noetherian. Hence R_H is noetherian.

Thus, by Lemma 2.6.10, the rewriting system (A, R_H) is complete. □

Returning to the proof of Theorem 3.1.1 (i) ($\Leftarrow$).

Proof of Theorem 3.1.1 (i) ($\Leftarrow$). By Lemma 3.4.4 the finite semigroup presentation $\langle A \mid R_H \rangle$ defines the subsemigroup H_{11}. By Lemma 3.4.5 the associated rewriting system (A, R_H) is complete. By definition of the Rees matrix semigroup S, the subsemigroup H_{11} is isomorphic to the group G. Therefore, $\langle A \mid R_H \rangle$ is a semigroup presentation which is a finite complete rewriting system and which defines the group G. □

Secondly we look to prove Theorem 3.1.1 (i) in the ($\Rightarrow$) direction.

Therefore our initial assumption is that the group G admits a presentation which is a finite complete rewriting system. Let G be defined by the semigroup presentation $\langle A \mid R \rangle$ which is a finite complete rewriting system. Let

$S = \mathcal{M}[G; I, \Lambda; P]$ be as defined in this Section 3.4. Then, by the definition of S, the subgroup H_{11} is isomorphic to G and therefore $\langle A \mid R \rangle$ is a presentation for H_{11} which is a finite complete rewriting system.

Next we define a proposed sparse generating set for S, with reference to Definition 3.3.2. Let the sets Q and Y be as defined in Lemma 3.3.1. We make a small change to the notation in the definition of Q as the set A is now our semigroup generating set for the group G. Therefore we have:

$$Q = \{(1, g, 1) : g \in A\} \text{ which is a semigroup generating set for } H_{11}.$$

In addition, using Notation 3.3.3, let
(i) A be the set of letters which correspond to the elements in the set Q and
(ii) B be the set of letters which correspond to the elements in the set Y.
Note that we are using the set A from the presentation $\langle A \mid R \rangle$ which we also used to define Q such that $\langle Q \rangle = H_{11}$. Again, we simplify notation and let $e_\lambda = (1, 1_g, \lambda) \in Y$.

Let $z \in A^+$ be a fixed word which represents the identity element of the group G. Let $R_S \subseteq (A \cup B)^+ \times (A \cup B)^+$ be the set of relations defined as:

$$R_S = R \cup \{(b_\lambda x, x) : b_\lambda \in B, x \in (A \cup B)\}$$

$$\cup \{(z b_\lambda, b_\lambda) : b_\lambda \in B\} \cup \{(b_\lambda z, z) : b_\lambda \in B\}.$$

Let η be the smallest congruence on $(A \cup B)^+$ which contains R_S.

Lemma 3.4.6. *Let $\langle A \cup B \mid R_S \rangle$ be the presentation as defined above. Then $\langle A \cup B \mid R_S \rangle$ is a presentation which defines the semigroup S and which is a finite complete rewriting system with respect to the sparse generating set $A \cup B$.*

Proof. It is immediate from the definitions that $\langle A \cup B \mid R_S \rangle$ is a finite presentation. It follows from Theorem 6.2 in [26] that $\langle A \cup B \mid R_S \rangle$ is a presentation defining S. It then follows from the proof of Theorem 3 and

Corollary 4 in [18] that $\langle A \cup B \mid R_S \rangle$ is a presentation which is a finite complete rewriting system. $\square$

Returning to the proof of Theorem 3.1.1 (i) ($\Rightarrow$).

Proof of Theorem 3.1.1 (i) ($\Rightarrow$). By Lemma 3.4.6 the presentation $\langle A \cup B \mid R_S \rangle$ is a semigroup presentation for S which is a finite complete rewriting system with respect to the sparse generating set $A \cup B$. $\square$

This completes the proof of Theorem 3.1.1 part (i).

3.5 New research regarding the special case of a single $\mathcal{L}$-class

Let us consider the special case for $S = \mathcal{M}[G; I, \Lambda; P]$ where $I = \{1, 2, \ldots, m\}$, $\Lambda = \{1\}$ with G and P as defined in Section 3.3. Our aim is to prove Theorem 3.1.1 part (ii). We first establish some basic definitions for this section.

As $\Lambda = \{1\}$ we make some adjustments to the definitions made in Section 3.3 which will simplify the notation and will apply for the remainder of this section.

Lemma 3.5.1. *Let $S = \mathcal{M}[G; I, \Lambda; P]$ be as defined in this section. Let the sets Q and Z be as defined in Lemma 3.3.1. Then $Q \cup Z$ is a finite generating set for S.*

Proof. The proof is contained within the proof of Lemma 3.3.1. $\square$

Next we specify our sparse generating set for this special case. Let the sets A and C be as defined in Section 3.3. The set B is empty and can be ignored. We call $A \cup C$ the sparse generating set for the semigroup S as defined in Definition 3.3.2.

Returning to the proof of Theorem 3.1.1 (ii)

Proof of Theorem 3.1.1 (ii). Given the above definitions that relate to the set up for this special case and as a consequence of the symmetry of the Rees matrix structure, our proof is the dual of the proofs in Section 3.4. □

3.6 Potential future work

3.6.1 Generalisation to Rees matrix semigroups over monoids

Conjecture 3.6.1. *Let $S = \mathcal{M}[M; I, \Lambda; P]$ be a Rees matrix semigroup where M is a monoid, the matrix P is normal and one of two conditions is true:*
(i) $I = \{1\}$ and $\Lambda = \{1, 2, \ldots, n\}$;
(ii) $I = \{1, \ldots, m\}$ and $\Lambda = \{1\}$.

Then the monoid M admits a presentation by a finite complete rewriting system if and only if S admits a presentation by a finite complete rewriting system with respect to some sparse generating set.

In this instance a sparse generating set $(A \cup B)$ or $(A \cup C)$ is defined in a similar manner to Lemma 3.3.1 but with M_A a semigroup generating set for the monoid M and other corresponding changes.

This result can be proved in much the same way as in Theorem 3.1.1 largely because the proofs for that theorem do not use the group property of inverse elements within G. One assumption has been added, namely that the matrix P is normal as this is not necessarily true in the monoid case.

In summary, the properties of the semigroup S that were assumed in order to prove Theorem 3.1.1 are as follows:

(i) That the rewriting system (A, R) is normalised. This remains true for the monoid case.

(ii) That the sparse generating set as defined could exist and would be sufficient to generate the semigroup. This remains true for the monoid case.

(iii) That the subgroup H_{11} is isomorphic to G. In the case of the monoid we do not have the same Green's structure with $\mathcal{H}$-classes isomorphic to G. Instead, the substructure we consider within the semigroup S is a submonoid which will be isomorphic to M. This submonoid must exist by definition of the Rees matrix structure of triples i.e. it is the submonoid generated by the set of elements $\{(1, m, 1) : m \in M_A\}$ where M_A is a finite generating set for the monoid M.

(iv) Each idempotent $e_{i\lambda}$ in Y and Z would be unique for each group of triples with $(1, m, \lambda)$ and $(i, m, 1)$ respectively, and would have the properties as used in the proof for Theorem 3.1.1. The relevant properties being that the idempotent is a left identity for $e_{1\lambda} \in Y$ or a right identity for $e_{i1} \in Z$.

In addition it would be necessary to check and possibly amend Theorem 3 and Corollary 4 in [18] to be applicable in the monoid case for the ($\Rightarrow$) direction. Note that Theorem 6.2 in [26] covers the monoid case.

3.6.2 Finitely many $\mathcal{R}$- and $\mathcal{L}$-classes

Extending Theorem 3.1.1 to the case
(iii) $I = \{1, \ldots, m\}$ and $\Lambda = \{1, \ldots, n\}$.

The proof in the ($\Leftarrow$) direction (passing from G to S) does hold in this more general case, since it follows from the proof of Theorem 3 in [18].

For the other direction ($\Rightarrow$) of the proof, initial investigations suggest that a proof will be possible with the additional condition that the semigroup S admits a presentation by a finite complete rewriting system which is *strongly minimal* (see following definition).

Definition 3.6.2. Let (A, R) be a rewriting system for a semigroup S which is complete. We say that the complete rewriting system R is *minimal* if, for each $(l, r) \in R$, r is irreducible with respect to R and l is irreducible with respect to $R - \{(l, r)\}$. We say that R is *strongly minimal* if it is minimal and each element of A is irreducible. See [12] for more details.

3.6.3 Sparse generating set

(i) It may be possible to remove the condition attached to Theorem 3.1.1 that we have a sparse generating set. This could be a natural property which is intrinsic to the problem. Alternatively, it may be possible to develop a method of finding this specific form of generating set, given the other properties and without affecting the fact that a presentation is complete.

(ii) Alternatively, the sparse generating set could be changed to select any element from the group $\mathcal{H}$-classes (that are not H_{11}), rather than an idempotent.

(iii) In the case where the Rees matrix semigroup does not naturally have a normalised matrix P, as it does for completely simple semigroups, then we could consider a more general constraint on P or even no constraint.

4

Finite derivation type and unitary subsemigroups with strict boundaries

4.1 Introduction

In this chapter we will investigate properties of a certain type of subsemigroup. Namely, we will consider subsemigroups which are left (respectively right) unitary with finite strict right (respectively left) boundary. In [16, Theorem 8.10] it was proved that such subsemigroups inherit from the semigroup the property of being finitely generated and finitely presented. A natural question in line with the theme of this research would be to ask whether the subsemigroup can also inherit the property of having a finite complete rewriting system. This problem remains out of reach. However, there is a related property of a semigroup having finite derivation type (FDT), which will be defined later in this chapter. A natural related question is whether FDT is inherited. The main result of this chapter will be to give a positive answer to this question. The main result of this chapter is:

Theorem 4.1.1. *Let S be a finitely presented semigroup with T a subsemigroup of S. Suppose S has finite derivation type. Then:*
(i) if T is left unitary and has finite strict right boundary in S then T also has finite derivation type;
(ii) if T is right unitary and has finite strict left boundary in S then T also has finite derivation type.

Before we proceed with the proof of this theorem we will first need to define the terms in the statement of the result. In addition, we will outline some results from [16] which will be needed for the proof of Theorem 4.1.1.

Note that in this chapter we are working with semigroups and subsemigroups. Therefore the presentations given are semigroup presentations and the rewriting systems are semigroup rewriting systems throughout the chapter, unless stated otherwise. In addition we are following the same conventions as those used in [20].

4.2 Introduction to boundary definitions and notation

The following definitions are taken from [16, Chapters 7 and 8] where more details can be found. Let S be a semigroup generated by a finite set A and defined by the semigroup presentation $\langle A \mid R\rangle$. Let T be a subsemigroup of S. Let η be the smallest congruence on A^+ which contains R. This notation S, $\langle A \mid R\rangle$, T and η will remain in force throughout this chapter.

At this point it is useful to recall the definition of $\mathcal{L}(A, T)$. For any subset of S, say (as in this case) the subset T, we set

$$\mathcal{L}(A, T) = \{w \in A^+ : w/\eta \in T\}$$

to be the set of all words in A^+ that represent elements of T in S. So we can see that $\mathcal{L}(A, T)$ is a subset of A^+ and in particular it can never contain the empty word.

The *right boundary of T in S with respect to A* is the set of elements of T that can be obtained by starting with an element in the complement $S \setminus T$ and right multiplying by a single generator from A, together with the elements of A that belong to T. In terms of the right Cayley graph of S, the boundary of T includes the terminal vertices of directed edges that start in $S \setminus T$ and end in T. In more detail, let $s \in S \setminus T$ and $sa_1 \in T$ with $a_1 \in A$. Then sa_1 is an element in the right boundary of T in S. If a belongs to both A and T then we also say that a belongs to the right boundary of T in S. The *left boundary* definition is dual but with the left and right perspectives reversed. In other words, let $s \in S \setminus T$ and $a_2 s \in T$ with $a_2 \in A$. Then $a_2 s$ is an element in the left boundary of S.

Definition 4.2.1. We use S^1 to denote S with an identity adjoined (even if it already has one), U to denote the complement $S \setminus T$ and U^1 to denote $S^1 \setminus T$. We use $\mathcal{B}_r(A,T)$, $\mathcal{B}_l(A,T)$ and $\mathcal{B}(A,T)$ to denote the right, left and two-sided boundaries respectively, of T in S. Then we have:

$$\mathcal{B}_l(A,T) = AU^1 \cap T = \{au : u \in U^1, a \in A\} \cap T$$

$$\mathcal{B}_r(A,T) = U^1 A \cap T = \{ua : u \in U^1, a \in A\} \cap T$$

$$\mathcal{B}(A,T) = \mathcal{B}_l(A,T) \cup \mathcal{B}_r(A,T).$$

We say T has *finite boundary in S* if for some finite generating set A of S the boundary $\mathcal{B}(A,T)$ is finite.

Clearly the sets defined above depend on the choice of generating set A. However, the finiteness (or otherwise) of these sets is independent of the choice of generating set (see Lemma 4.2.2). Thus we may speak of T being a subsemigroup with finite (left, right or two-sided) boundary without reference to the generating set for S.

Lemma 4.2.2. *[16, Proposition 7.3] Let S be a finitely generated semigroup, let T be a subsemigroup of S and let A and B be two finite generating sets for S. Then $\mathcal{B}_r(A,T)$ is finite if and only if $\mathcal{B}_r(B,T)$ is finite. Also, $\mathcal{B}_l(A,T)$ is finite if and only if $\mathcal{B}_l(B,T)$ is finite.*

Definition 4.2.3. We call $w \in A^+$ a *strict left boundary word* of T in S with respect to A if $w \in \mathcal{L}(A,T)$ and no proper suffix of w belongs to $\mathcal{L}(A,T)$. Similarly, we call w a *strict right boundary word* of T in S with respect to A if $w \in \mathcal{L}(A,T)$ and no proper prefix of w belongs to $\mathcal{L}(A,T)$. We denote the set of strict right and left boundary words, by $\mathcal{SWB}_r(A,T)$ and $\mathcal{SWB}_l(A,T)$ respectively.

Definition 4.2.4. The *strict left boundary* of T in S with respect to A is $\mathcal{SB}_l(A,T) = \mathcal{SWB}_l(A,T)/\eta$ and the *strict right boundary* of T in S with respect to A is $\mathcal{SB}_r(A,T) = \mathcal{SWB}_r(A,T)/\eta$. Also, the *strict two-sided boundary* is defined as $\mathcal{SB}(A,T) = \mathcal{SB}_l(A,T) \cup \mathcal{SB}_r(A,T)$.

Note that $\mathcal{SB}_l(A,T) \subseteq \mathcal{B}_l(A,T)$ and $\mathcal{SB}_r(A,T) \subseteq \mathcal{B}_r(A,T)$. We say that the strict left, right or two-sided boundary of T in S is finite if for some finite generating set A of S, the sets $\mathcal{SB}_l(A,T)$, $\mathcal{SB}_r(A,T)$, $\mathcal{SB}(A,T)$ are respectively finite. The question of whether the strict right, left and two-sided boundaries are finite, depends on the choice of generating set. This is illustrated in the following example.

Example 4.2.5. The following is an example from [16, Example 8.3].
We use $\mathbb{Z}_2$ to denote the set $\{0,1\}$ with the operation of addition modulo 2, which is the cyclic group of order 2 written additively. Let $S = \mathbb{Z} \oplus \mathbb{Z}_2$ which is generated, as a semigroup, by $A = \{(1,0),(-1,0),(0,1)\}$ and also by the set $B = \{(1,1),(-1,1),(0,1)\}$. Let $T = \mathbb{Z} \oplus \{0\}$, which is a subsemigroup of S. Then we claim that $\mathcal{SB}(A,T)$ is infinite while $\mathcal{SB}(B,T)$ is finite.

For all $b_1, b_2 \in B$ we have $b_1 + b_2 \in T$ and so the strict (left, right and two-sided) boundaries all equal $\{(0,0),(-2,0),(2,0),(1,0),(-1,0)\}$.

To see that $\mathcal{SB}(A,T)$ is infinite note that:

$$(0,1) + \underbrace{\big((1,0) + (1,0) + \ldots + (1,0)\big)}_{m\text{ - times}} + (0,1) = (m,0)$$

belongs to $\mathcal{SB}(A,T)$ for all $m \in \mathbb{N}$, since for all k

$$(0,1) + \underbrace{(1,0) + (1,0) + \ldots + (1,0)}_{k\text{ - times}} = (k,1) \notin T.$$

4.3 Introduction to unitary subsemigroups with strict boundaries

Definition 4.3.1. [25, Page 63 Exercise 20] Let T be a subsemigroup of the semigroup S. Then T is *right unitary* if

$$\forall s \in S, \quad \forall t \in T, \quad st \in T \Leftrightarrow s \in T;$$

and *left unitary* if

$$\forall s \in S, \quad \forall t \in T, \quad ts \in T \Leftrightarrow s \in T;$$

and *unitary* if it is both left and right unitary.

Example 4.3.2. Two examples of such subsemigroups:
(a) Subgroups of groups are always left and right unitary subsemigroups.
(b) Let M be a monoid and let G be the group of units of M. Then G is both a left and right unitary semigroup of M.

These conditions on a semigroup and its subsemigroup relate closely to the questions considered in Chapter 3 with respect to Rees matrix semigroups. If we recall the case of a Rees matrix semigroup S which has a single $\mathcal{R}$-class and finitely many $\mathcal{L}$-classes where we consider a subgroup H which is isomorphic to a group $\mathcal{H}$-class, namely H_{11}. Then, with respect to the sparse generating set (see Chapter 3), the subgroup H has a finite strict right boundary in S. In addition, the subgroup H is a left unitary subsemigroup. So we have an example of a subgroup which is a left unitary subsemigroup and has a finite strict right boundary. Similarly, recall the dual case where we have a Rees matrix semigroup S which has a single $\mathcal{L}$-class and finitely many $\mathcal{R}$-classes where we consider a subgroup H which is isomorphic to a group $\mathcal{H}$-class. Here we have an example of a subgroup H which is a right unitary subsemigroup and has a finite strict left boundary (with respect to the sparse generating set) in S.

This leads to a related open question regarding the inheritance of being a finite complete rewriting system in the case of a subsemigroup which is left (resp. right) unitary and has a finite strict right (resp. left) boundary. This links to the Rees matrix semigroup open question (see Section 3.1) regarding the inheritance of being a finite complete rewriting system by a subsemigroup (i.e. an $\mathcal{H}$-class).

4.4 Introduction to Finite Derivation Type (FDT)

4.4.1 Introduction

The property of having finite derivation type was introduced for monoids by C.C. Squier, F. Otto and Y. Kobayashi in [51]. It is a finiteness condition which is satisfied by every monoid which has a FCRS.

The properties of having finite derivation type (FDT) and being a finite complete rewriting system (FCRS) are closely related for a semigroup. If a finitely presented semigroup has a FCRS, then that semigroup also has FDT. Furthermore, the property FDT is independent of the choice of presentation for the semigroup. In contrast, the property of being a FCRS is dependent on the choice of presentation, which is one of the things which makes it a difficult property to work with. So considering the question of whether a semigroup has FDT can be helpful in understanding whether it may or may not have a FCRS. Specifically, if it can be proven that a finitely presented semigroup does not have FDT, then there does not exist a presentation for the semigroup that is a FCRS. This is a result of the following theorems:

Theorem 4.4.1. *[51, Theorem 4.3] If two finite presentations $\mathcal{P}_1$ and $\mathcal{P}_2$ define the same monoid (i.e. they are Tietze equivalent) then $\mathcal{P}_1$ has FDT if and only if $\mathcal{P}_2$ has FDT.*

Theorem 4.4.2. *[51, Theorem 5.3] Let M be a finitely presented monoid. If M admits a presentation by a finite complete rewriting system then M has FDT.*

Clearly, the above theorems are specific to monoids and monoid presentations.

However, the following theorem and definition enable us to use corresponding theorems which are with respect to semigroups and semigroup presentations.

Theorem 4.4.3. *[38, Theorem 3] Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be finite semigroup presentations defining the same semigroup. The presentation $\mathcal{P}_1$ has FDT if and only if $\mathcal{P}_2$ also has FDT.*

Definition 4.4.4. [19, Definition 4] A finitely presented semigroup S has finite derivation type if some (and hence any by [51, Theorem 4.3] and [38, Theorem 3]) finite presentation for S has finite derivation type.

4.4.2 Definitions

The following definitions are courtesy of [19] and [20] where more detail may be found.

Let $\mathcal{P} = \langle A \mid R \rangle$ be a presentation for the semigroup S. The *derivation graph* associated with $\mathcal{P}$ is an infinite graph $\Gamma = \Gamma(\mathcal{P}) = (V, E, \iota, \tau, ^{-1})$ with *vertex set* $V = A^+$, and *edge set* E consisting of the following collection of 4-tuples:

$$\{(w_1, r, \epsilon, w_2) : w_1, w_2 \in A^*,\ r = (r_{+1}, r_{-1}) \in R,\ \epsilon \in \{+1, -1\}\}.$$

The functions $\iota, \tau : E \to V$ associate with each edge $\mathbb{E} = (w_1, r, \epsilon, w_2)$ (with $r = (r_{+1}, r_{-1}) \in R$) its initial and terminal vertices $\iota\mathbb{E} \equiv w_1 r_\epsilon w_2$ and $\tau\mathbb{E} \equiv w_1 r_{-\epsilon} w_2$, respectively. The mapping $^{-1} : E \to E$ associates with each edge $\mathbb{E} = (w_1, r, \epsilon, w_2)$ an inverse edge $\mathbb{E}^{-1} = (w_1, r, -\epsilon, w_2)$.

A *path* is a sequence of edges $\mathbb{P} = \mathbb{E}_1 \circ \mathbb{E}_2 \circ \ldots \circ \mathbb{E}_n$ where $\tau\mathbb{E}_i \equiv \iota\mathbb{E}_{i+1}$ for $i = 1, \ldots, n-1$. Here $\mathbb{P}$ is a path from $\iota\mathbb{E}_1$ to $\tau\mathbb{E}_n$ and we extend the mappings ι and τ to paths by defining $\iota\mathbb{P} \equiv \iota\mathbb{E}_1$ and $\tau\mathbb{P} \equiv \tau\mathbb{E}_n$. The *length of a path* is the number of edges in the path. In the previous example, the length of $\mathbb{P}$ is n. The *width of a path* is the maximum length of the vertex words in the edges of the path i.e. *max* $(|\iota\mathbb{E}_i|, |\tau\mathbb{E}_n|)$ with $i \in \{1, \ldots, n\}$. The inverse of a path $\mathbb{P} = \mathbb{E}_1 \circ \mathbb{E}_2 \circ \ldots \circ \mathbb{E}_n$ is the path $\mathbb{P}^{-1} = \mathbb{E}_n^{-1} \circ \mathbb{E}_{n-1}^{-1} \circ \ldots \circ \mathbb{E}_1^{-1}$, which is a path from $\tau\mathbb{P}$ to $\iota\mathbb{P}$. A *closed path* is a path $\mathbb{P}$ satisfying $\iota\mathbb{P} \equiv \tau\mathbb{P}$. For two paths $\mathbb{P}$ and $\mathbb{Q}$ with $\tau\mathbb{P} \equiv \iota\mathbb{Q}$ the composition $\mathbb{P} \circ \mathbb{Q}$ is defined. To unclutter the notation we will omit the symbol $\circ$ when composing paths, writing simply $\mathbb{P}\mathbb{Q}$ in place of $\mathbb{P} \circ \mathbb{Q}$.

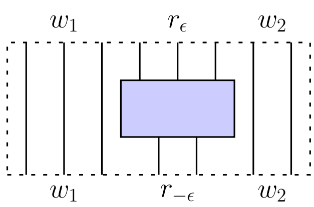
We denote the set of paths in Γ by $P(\Gamma)$, where for each vertex $w \in V$ we include a path 1_w with no edges, called the *empty path* at w. We call a path $\mathbb{P}$ *positive* if it is either empty or it contains only edges of the form $(w_1, r, +1, w_2)$. We use $P_+(\Gamma)$ to denote the set of all *positive paths* in Γ. Dually we have the notion of a *negative path*, and $P_-(\Gamma)$ denotes the set of all negative paths. The free monoid A^* acts on both sides of the set of edges E of Γ, where if $\mathbb{E} = (w_1, r, \epsilon, w_2)$ and $x, y \in A^*$ we have:

$$x \cdot \mathbb{E} \cdot y = (xw_1, r, \epsilon, w_2 y).$$

This extends naturally to a two-sided action of A^* on $P(\Gamma)$ where for a path $\mathbb{P} = \mathbb{E}_1 \circ \mathbb{E}_2 \circ \ldots \circ \mathbb{E}_n$ we define

$$x \cdot \mathbb{P} \cdot y = (x \cdot \mathbb{E}_1 \cdot y) \circ (x \cdot \mathbb{E}_2 \cdot y) \circ \ldots \circ (x \cdot \mathbb{E}_n \cdot y).$$

The edge $\mathbb{E} = (w_1, r, \epsilon, w_2)$ can be represented geometrically by an object called a *picture* as follows:



These are also called *atomic pictures* and are always directed downwards in the sense that $\iota\mathbb{E}$ is the word read along the top of the picture and $\tau\mathbb{E}$ is the word read along the bottom. The rectangle in the centre of the picture is called a *transistor* and corresponds to the relation r from the presentation, while the line segments in the diagram are called *wires* (or *strings*) with each wire labelled by a unique letter from the free monoid A^*. The monoid picture for the inverse $\mathbb{E}^{-1}$

of an edge is obtained by taking the vertical mirror image of the picture of $\mathbb{E}$. By "stacking" such pictures one on top of the other, and adjoining corresponding wires, we obtain pictures for arbitrary paths of the graph $\Gamma(\mathcal{P})$. A picture of a path $\mathbb{P}$ with $\iota\mathbb{P} \equiv \tau\mathbb{P}$ is called *closed* or *spherical*.

For every $r \in R$ and $\epsilon = \pm 1$ define $A_r^\epsilon = (1, r, \epsilon, 1)$. Note that here we are using 1 to represent the empty word as ϵ is used to define whether the edge is positive or negative. We call such edges *elementary*. As we have an elementary edge for each relation in R then every edge of $\Gamma(\mathcal{P})$ can be written uniquely in the form $\alpha \cdot \mathbb{A} \cdot \beta$ where $\alpha, \beta \in A^*$ and $\mathbb{A}$ is elementary. Furthermore, as an edge in $\Gamma(\mathcal{P})$ corresponds to a single application of a rewriting rule, then for an arbitrary path $\mathbb{P}$, $\iota\mathbb{P}$ and $\tau\mathbb{P}$ represent the same element of the semigroup defined by the presentation. Thus $\Gamma(\mathcal{P})$ is called the *derivation graph* of the presentation. It follows that there is a one-to-one correspondence between the elements of the semigroup S and the connected components of the derivation graph $\Gamma(\mathcal{P})$.

If $\mathbb{P}$ and $\mathbb{Q}$ are paths such that $\iota\mathbb{P} \equiv \iota\mathbb{Q}$ and $\tau\mathbb{P} \equiv \tau\mathbb{Q}$ then we say that $\mathbb{P}$ and $\mathbb{Q}$ are *parallel*, and write $\mathbb{P} \parallel \mathbb{Q}$. We use $\parallel$ to denote the subset of $P(\Gamma) \times P(\Gamma)$ which comprises all pairs of parallel paths.

For any two edges $\mathbb{E}_1, \mathbb{E}_2$ in the graph we define the following path (which is illustrated in Figure 4.4.1 below):

$$[\mathbb{E}_1, \mathbb{E}_2] = (\mathbb{E}_1 \cdot \iota\mathbb{E}_2) \circ (\tau\mathbb{E}_1 \cdot \mathbb{E}_2) \circ (\mathbb{E}_1^{-1} \cdot \tau\mathbb{E}_2) \circ (\iota\mathbb{E}_1 \cdot \mathbb{E}_2^{-1}).$$

Note that it is an immediate consequence of this definition that:

$$\alpha \cdot [\mathbb{E}_1, \mathbb{E}_2] \cdot \beta = [\alpha \cdot \mathbb{E}_1, \mathbb{E}_2 \cdot \beta] \text{ with } \alpha, \beta \in A^*.$$

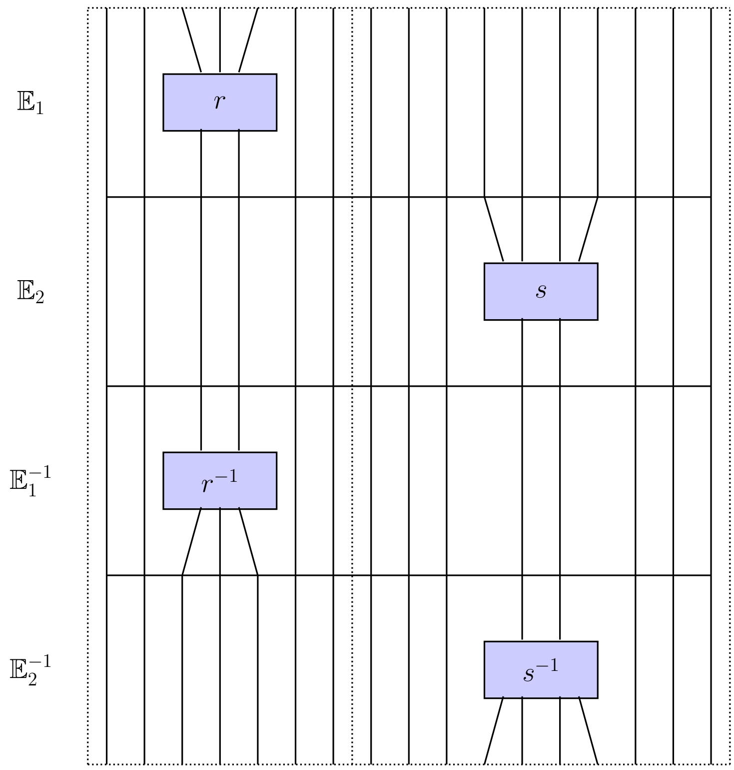


Figure 4.4.1: Path for $[\mathbb{E}_1, \mathbb{E}_2]$

The following definition was introduced in [51, Definition 3.2] and has been updated to match the notation used in this chapter.

An equivalence relation $\sim$ on $P(\Gamma)$ is called a *homotopy relation* if it is contained in $\|$ and satisfies the following conditions:

(H1) If $\mathbb{E}_1$ and $\mathbb{E}_2$ are edges of Γ, then

$$(\mathbb{E}_1 \cdot \iota\mathbb{E}_2)(\tau\mathbb{E}_1 \cdot \mathbb{E}_2) \sim (\iota\mathbb{E}_1 \cdot \mathbb{E}_2)(\mathbb{E}_1 \cdot \tau\mathbb{E}_2)$$

(H2) For any $\mathbb{P}, \mathbb{Q} \in P(\Gamma)$ and $x, y \in A^*$

$$\mathbb{P} \sim \mathbb{Q} \Rightarrow x \cdot \mathbb{P} \cdot y \sim x \cdot \mathbb{Q} \cdot y.$$

(H3) For any $\mathbb{P}, \mathbb{Q}, \mathbb{R}, \mathbb{S} \in P(\Gamma)$ with $\tau\mathbb{R} \equiv \iota\mathbb{P} \equiv \iota\mathbb{Q}$ and $\iota\mathbb{S} \equiv \tau\mathbb{P} \equiv \tau\mathbb{Q}$

$$\mathbb{P} \sim \mathbb{Q} \Rightarrow \mathbb{RPS} \sim \mathbb{RQS}.$$

(H4) If $\mathbb{P} \in P(\Gamma)$ then $\mathbb{P}\mathbb{P}^{-1} \sim 1_{\iota\mathbb{P}}$, where $1_{\iota\mathbb{P}}$ denotes the empty path at the vertex $\iota\mathbb{P}$.

The collection of all homotopy relations is closed under arbitrary intersection and $\|$ is a homotopy relation. Therefore, for any subset C of $\|$ there is a unique smallest homotopy relation $\sim_C$ on $P(\Gamma)$ containing C. We call this the *homotopy relation generated by* C. A subset C of $\|$ that generates $\|$ is called a *homotopy base for* Γ.

A presentation $P = \langle A \mid R \rangle$ has *finite derivation type* (or FDT for short) if there is a finite homotopy base for $\Gamma = \Gamma(P)$. A finitely presented semigroup S has finite derivation type if some finite presentation for S has finite derivation type.

It can be shown using (H1) to (H4) that a set B of parallel paths is a homotopy base if and only if the set $\{(\mathbb{P} \circ \mathbb{Q}^{-1}, 1_{\iota\mathbb{P}}) : (\mathbb{P}, \mathbb{Q}) \in B\}$ is. Hence we say that a set C of closed paths is a homotopy base if $\{(\mathbb{P}, 1_{\iota\mathbb{P}}) : \mathbb{P} \in C\}$is a homotopy base. So a homotopy base X for $\Gamma = \Gamma(P)$ may be given either as a subset of $\|$, so that X is a set of $\|$ paths, or may be given as a set of closed paths. Sometimes X is referred to as a homotopy base of a presentation P, rather than the graph $\Gamma(P)$. Also, when it is clear from context, as a homotopy base of a semigroup.

When considering FDT for monoids, it makes no difference whether we work with semigroup presentations or with monoid presentations. Note that in this chapter we will work with semigroup presentations.

4.5 An existing proof for presentations of unitary subsemigroups with strict boundaries

The new research in this chapter builds on the results contained in [16, Chapter 8]. Specific theorems from this work and the presentations defined in order to prove them, provide a starting point for the proof of the main new result in this chapter. As such they are reproduced here with relevant details.

In addition to stating some definitions and results from [16, Chapter 8], we will also include proofs of a few of the results (taken from [16]). In each case we have included these proofs here because later in the chapter when our new results on FDT are included, the proof will use similar (but more complicated) ideas. So by including these proofs from [16] we hope that it will make the generalisations that come later in the chapter easier for the reader to follow.

The main theorem which is of interest is as follows:

Theorem 4.5.1. *[16, Theorem 8.10] Let S be a finitely presented semigroup with T a subsemigroup of S. Then:*

(i) if T is left unitary and has finite strict right boundary in S then T is finitely presented;

(ii) if T is right unitary and has finite strict left boundary in S then T is finitely presented.

The proof of this theorem uses the Reidemeister-Schreier type rewriting process and Theorem 2.9.1, see Subsection 2.9 for more details. We will look at the proof for part (i). Note that the definitions in this section will apply for the remainder of this chapter unless specified to the contrary.

Let $\langle A \mid R \rangle$ be a finite presentation for the semigroup S and T be a left unitary subsemigroup of S with finite strict right boundary. Let η be the smallest congruence on A^+ containing R.

First a suitable finite generating set for the subsemigroup T is required. The following theorem and corollary assist with this. Note that the proof of this theorem contains a method of reading an element which belongs to $\mathcal{L}(A, T)$ from left to right and decomposing the word into principal factors. This process will be used later.

Theorem 4.5.2. *[16, Theorem 8.7] Let S be a semigroup generated by the finite set A. Let T be a left unitary subsemigroup of S. Then $\langle \mathcal{SB}_r(A, T) \rangle = T$.*

Proof. Let t be an arbitrary element in T and write $t = a_1a_2\dots a_n$ where each a_i is in A. Then we read the element from left to right, that is we consider a_1 then a_1a_2 and so on. We are looking to find the shortest non-empty prefix of t which is in T. This will decompose the element $a_1a_2\dots a_n$ into uv where u is on the strict right boundary of T and v is in T since uv is in T, and T is left unitary. Then we continue the process by reading the element v from left to right. In this way we decompose the element into *principal factors* which all belong to the strict right boundary of T in S. Hence we can see that $\mathcal{SB}_r(A,T)$ generates T. $\square$

Corollary 4.5.3. *[16, Corollary 8.8] Let S be a finitely generated semigroup and let T be a subsemigroup of S. Then:*

(i) if T is left unitary and has a finite strict right boundary in S then T is finitely generated;

(ii) if T is right unitary and has a finite strict left boundary in S then T is finitely generated.

It is not a general fact that subsemigroups with finite strict boundaries are finitely generated. It is the additional condition that the subsemigroup is left (or right) unitary which makes the right (or left) strict boundary a finite generating set.

Recall that $\mathcal{SB}_r(A,T)$ is a set of elements in T. We now look to find a unique set of words to represent each of these elements and so create a generating set of words. Define $\theta : A^+ \to S$ where if $w \in A^+$ then $w\theta \mapsto (w/\eta)$. So the homomorphism θ maps words in A^+ to the element in S that they represent.

We choose and fix a transversal $\mathcal{R}$ of the η-classes of $\mathcal{L}(A,T)$ with the following property:

- For every $w \in \mathcal{R}$, if $w/\eta \in \mathcal{SB}_r(A,T)$ then $w \in \mathcal{SWB}_r(A,T)$

i.e. representatives of strict right boundary elements are always chosen to be strict right boundary words. For every $w \in \mathcal{L}(A,T)$ define $\bar{w} = (w/\eta) \cap \mathcal{R}$, the fixed chosen word in $\mathcal{R}$ that equals w in S.

Let

$$Y = \{v : v \in \mathcal{R} \cap \mathcal{L}(A, \mathcal{SB}_r(A,T))\}$$

which is a set of strict right boundary words which represent each element in the strict right boundary $\mathcal{SB}_r(A,T)$ of T in S. Hence Y is a finite generating set for T, by Theorem 4.5.2 and as $\mathcal{SB}_r(A,T)$ is finite (by definition of T).

Next we define a new alphabet B in one to one correspondence with the generating words in the generating set Y as follows:

$$B = \{b_v : v \in \mathcal{R} \cap \mathcal{L}(A, \mathcal{SB}_r(A,T))\}.$$

By taking words from $\mathcal{R}$ we have a finite generating set in terms of words whereas $\mathcal{SWB}_r(A,T)$ could be an infinite set of words. Now we can define a representation mapping, let $\psi : B^+ \to A^+$ be the unique homomorphism which maps $b_v \mapsto v$.

In addition we define a rewriting mapping $\phi : \mathcal{L}(A,T) \to B^+$ so that we can map all of the words in $\mathcal{L}(A,T)$ to the new alphabet B. Let $w \in \mathcal{L}(A,T)$ and set $w \equiv \alpha\beta$ where α is the shortest non-empty prefix of w that belongs to $\mathcal{L}(A,T)$. The key fact that simplifies our inspection of w is that T is left unitary and by that definition we know that $\beta \in \mathcal{L}(A,T)$. So we can work along our word w from left to right, taking subsets of letters (or even single letters) that represent elements in T. So now we can define ϕ inductively as follows:

$$w\phi = (\alpha\beta)\phi = \begin{cases} b_{\bar{\alpha}} \text{ if } \beta \text{ is the empty word} \\ \\ b_{\bar{\alpha}}(\beta\phi) \text{ otherwise.} \end{cases}$$

We use $\bar{\alpha}$, rather than α, so that we have a mapping to a unique word in the same η-class, representing the same element in T as α. This then matches our finite generating set Y and so our mapped word $w\phi$ is composed from our alphabet B.

Note that if $w \in \mathcal{L}(A,T)$ then $(w\phi)\psi = w$ in S but $(w\phi)\psi \not\equiv w$ so we do not get back to the identical word but to a word composed of our unique fixed right boundary words and which represents the same element.

Now we are in a position to use Theorem 2.9.1 and say that the subsemigroup T is defined by the presentation with generators B and relations

$$b_v = v\phi \tag{4.5.1}$$

$$(w_1w_2)\phi = (w_1\phi)(w_2\phi) \tag{4.5.2}$$

$$(w_3xw_4)\phi = (w_3yw_4)\phi \tag{4.5.3}$$

where $v \in \mathcal{R} \cap \mathcal{L}(A, \mathcal{SB}_r(A,T))$, $w_1, w_2 \in \mathcal{L}(A,T)$, $w_3, w_4 \in A^*$, $(x = y) \in R$ and $w_3xw_4 \in \mathcal{L}(A,T)$.

As we have seen before, this gives us an infinite set of relations and so we now look for a finite subset which is equivalent. Firstly, we can see that all of the relations (4.5.1) are trivial, that is they each have the form $b_v \equiv v\phi$, as per Lemma 4.5.4 below.

Lemma 4.5.4. *The relations (4.5.1) are trivial i.e.* $b_v \equiv v\phi$.

Proof. Let $v \in \mathcal{R} \cap \mathcal{L}(A, \mathcal{SB}_r(A,T))$.
Now $v \in \mathcal{L}(A, \mathcal{SB}_r(A,T)) \Rightarrow v/\eta \in \mathcal{SB}_r(A,T)$.
Now $v \in \mathcal{R}$ and $v/\eta \in \mathcal{SB}_r(A,T) \Rightarrow v \in \mathcal{SWB}_r(A,T)$ by definition of $\mathcal{R}$.

Since the shortest non-empty prefix of v that is in $\mathcal{L}(A,T)$ is v itself (by definition of $\mathcal{SWB}_r(A,T)$) it follows that, by definition of ϕ, we have $(v)\phi \equiv b_{\bar{v}}$.

But $v \in \mathcal{R} \Rightarrow \bar{v} \equiv v$ hence $(v)\phi \equiv b_{\bar{v}} \equiv b_v$. □

Next we aim to show that relations (4.5.2) are also all trivial.

Lemma 4.5.5. *[16, Lemma 8.11] Let* $w \in \mathcal{L}(A,T)$. *Then* w *may be written* *uniquely as* $w \equiv \alpha_1 \dots \alpha_k$ *where* $k \geq 1$ *and* $\alpha_i \in \mathcal{SWB}_r(A,T)$ *for all* i*. Moreover, this decomposition satisfies:*

$$w\phi \equiv (\alpha_1 \dots \alpha_k)\phi \equiv (\alpha_1\phi) \dots (\alpha_k\phi) \equiv b_{\bar{\alpha_1}} \dots b_{\bar{\alpha_k}}.$$

Also, given $w_1, w_2 \in \mathcal{L}(A,T)$*, where* w_1 *and* w_2 *decompose as* $\alpha_1 \dots \alpha_k$ *and* $\beta_1 \dots \beta_l$ *respectively, we have:*

$$(w_1w_2)\phi \equiv (\alpha_1 \dots \alpha_k\beta_1 \dots \beta_l)\phi \equiv b_{\bar{\alpha_1}} \dots b_{\bar{\alpha_k}} b_{\bar{\beta_1}} \dots b_{\bar{\beta_l}} \equiv (w_1\phi)(w_2\phi).$$

Proof. Summarising the original proof.
See the proof of Theorem 4.5.2 for details of decomposition into principal factors α_i and then the first statement follows from the definition of ϕ. For the second part it follows that $\alpha_i \in \mathcal{SB}_r(A,T)$ and $\beta_i \in \mathcal{SB}_r(A,T)$ for all i. Then from the definition of ϕ we have:

$$(\alpha_1 \dots \alpha_k\beta_1 \dots \beta_l)\phi \equiv (\alpha_1\phi) \dots (\alpha_k\phi)(\beta_1\phi) \dots (\beta_l\phi).$$

□

Now we claim the following is a presentation for T (which leaves the lemma that follows to be proved):

Theorem 4.5.6. *[16, Theorem 8.12] Let* S *be a finitely generated semigroup defined by a presentation* $\langle A \mid R \rangle$ *with* A *finite. Let* T *be a left unitary semigroup of* S *with finite strict boundary. Then with the above notation,* $\langle B \mid U \rangle$ *is a presentation for* T *where*

$$U = \{\ (u,v) \in B^+ \times B^+ : u\psi = v\psi \text{ in } S \text{ and } |uv| \leq max\{|\alpha\beta| : (\alpha = \beta) \in R\}\ \}.$$

In particular, if R *is finite then* T *is finitely presented.*

So the previous theorem says that if $\langle A \mid R \rangle$ is a finite presentation for S, then we obtain a finite presentation for T by taking a presentation $\langle B \mid U \rangle$ where the generators B correspond to the elements of the strict right boundary of T in S, and U consists of all the relations that hold in T between the generators B, where the length of the relation is not longer than the maximum of the lengths of the relations that appear in R.

Lemma 4.5.7. *[16, Lemma 8.13] The relations $(w\alpha v)\phi = (w\beta v)\phi$ where $w, v \in A^*$, $(\alpha = \beta) \in R$ and $w\alpha v \in \mathcal{L}(A, T)$ are consequences of U.*

In summary, the condition that the subsemigroup is left unitary means that decomposition of words from left to right behaves in a more predictable and helpful way. This is essential to the proofs and works together with there being a finite strict right boundary.

The dual result can be proved by decomposing words into principal factors from right to left rather than from left to right.

4.6 New research relating FDT and unitary subsemigroups with strict boundaries

In this section we will prove the main result in this chapter, which is:

Theorem 4.1.1. *Let S be a finitely presented semigroup with T a subsemigroup of S. Suppose S has finite derivation type. Then:*
(i) if T is left unitary and has finite strict right boundary in S then T also has finite derivation type;
(ii) if T is right unitary and has finite strict left boundary in S then T also has finite derivation type.

4.6.1 Defining basic conditions

Let S be a finitely presented semigroup with presentation $\langle A \mid R \rangle$. Let T be a left unitary subsemigroup of S with finite strict right boundary in S with respect to A. Denote the presentation which defines S as $\mathcal{P} = \langle A \mid R \rangle$. Suppose that the semigroup S has finite derivation type which, by Theorem 4.4.3, means that the presentation $\mathcal{P}$ has finite derivation type.

In the previous subsection Theorem 4.5.6 gives a presentation for T. In this subsection we will take that same presentation for T, using the same notation, and we will construct a homotopy base for T with respect to that presentation. To that end, we first recall some definitions from the previous Section 4.5.

Let η and $\theta : A^+ \to S$ be as defined in Section 4.5. So the strict right boundary of T is equal to $\mathcal{SB}_r(A,T) = \mathcal{SWB}_r(A,T)\theta \subseteq T \subseteq S$. Recall the transversal $\mathcal{R}$ of the η-classes of $\mathcal{L}(A,T)$ and that if $w \in \mathcal{L}(A,T)$ we have $\bar{w} = (w/\eta) \cap \mathcal{R}$, the fixed chosen word in $\mathcal{R}$ that equals w in S, all as per Section 4.5. Note that the representatives of strict right boundary elements in the transversal $\mathcal{R}$ are always chosen to be strict right boundary words.

Let the finite generating set Y and the new alphabet B be as defined in Section 4.5. Let $\phi : \mathcal{L}(A,T) \to B^+$ and $\psi : B^+ \to A^+$ be the rewriting and representation mappings defined in Section 4.5. We start with the following lemma which is a consequence of the definitions of ϕ and ψ.

Lemma 4.6.1. *For every $w \in B^+$ we have $(w\psi)\phi \equiv w$.*

Proof. Let $w \equiv b_{v_1} b_{v_2} \dots b_{v_m}$ be an arbitrary word in B^+. Then by definition of ψ we have $w\psi \equiv v_1 v_2 \dots v_m \in \mathcal{L}(A,T)$. Recall $v_1, \dots, v_m \in Y$, the generating set for T, and that these are strict right boundary words, chosen to uniquely represent the elements in the strict right boundary $\mathcal{SB}_r(A,T)$ of T in S. See Section 4.6.1 for the precise definition of the finite generating set Y for the subsemigroup T and the homomorphism ϕ. Now consider $(w\psi)\phi$ and recall that by the proof of Theorem 4.5.2 every $t \in \mathcal{L}(A,T)$ has a unique decomposition into words that belong to the strict right boundary of T. So when we decompose $w\psi \equiv v_1 v_2 \dots v_m$ the principal factors will be $v_1, v_2, \dots, v_m \in Y$. Then by applying ϕ to get $(w\psi)\phi$ we will have $b_{\bar{v}_i} \equiv b_{v_i}$ for every $i \in \{1, \dots, m\}$. So now we have

$$(w\psi)\phi \equiv (b_{v_1} b_{v_2} \dots b_{v_m})\psi\phi \equiv (v_1 v_2 \dots v_m)\phi \equiv b_{v_1} b_{v_2} \dots b_{v_m} \equiv w.$$

□

Let the set of relations U in $B^+ \times B^+$ be as defined in Theorem 4.5.6. Recall that by Theorem 4.5.6 the semigroup T is defined by the finite presentation $\langle B \mid U \rangle$.

Recall that by Theorem 4.4.3, if FDT is a property for one presentation it is true for all presentations for the same semigroup. Thus we can now look to prove that the presentation $\mathcal{Q} = \langle B \mid U \rangle$ has property finite derivation type.

4.6.2 Outline of the proof of Theorem 4.1.1

In Section 4.5 an infinite presentation is created for T with a finite generating set and an infinite set of defining relations and then it is proved that there is a finite subset of the defining relations of which all the other defining relations are a consequence. Similarly for FDT we will create an infinite homotopy base before finding a finite subset which will give us a finite homotopy base for T and thus prove that T has FDT. We will make use of the fact that words in T decompose into principal factors in a certain way which will simplify the special cases that need to be considered. The proof will be given for a left unitary subsemigroup with finite strict right boundary in S. The proof for a right unitary subsemigroup with finite strict left boundary is the dual of this with left and right reversed. The sections below include definitions and lemmas to enable the following steps in the proof to be completed:

(i) Extend ϕ and ψ to map between a certain subset of the vertices of the derivation graph for $\Gamma(\mathcal{P})$, namely the set $\mathcal{L}(A, T)$, and the derivation graph $\Gamma(\mathcal{Q})$ by defining maps for edges and paths. Note that vertices map as for words and for ϕ we are only mapping from the words in $\mathcal{L}(A, T)$ which is a strict subset of A^+.

(ii) Define an infinite homotopy base for $\Gamma(\mathcal{Q})$ using general results as used in [19, Lemma 9].

(iii) Identify a finite subset of the infinite homotopy base as defined in (ii).

(iv) Prove that all paths in our infinite homotopy base are a consequence of those in the finite set (iii).

4.6.3 Extending the definitions of ϕ and ψ to derivation graphs

In this section we will extend the definitions for ϕ and ψ so that we can map between a certain subset of the edges and paths of the derivation graph $\Gamma(\mathcal{P})$ for the semigroup presentation $\langle A \mid R\rangle$ and the edges and paths of the derivation graph $\Gamma(\mathcal{Q})$ for the subsemigroup presentation $\langle B \mid U\rangle$.

At this point it is useful to recall the domains and ranges of the mappings ϕ and ψ. The function ψ maps B^+ to A^+. However, ϕ does not map A^+ to B^+ as ϕ is only defined on certain words, specifically ϕ maps $\mathcal{L}(A, T)$ to B^+. So ϕ maps a certain subset of the vertices of $\Gamma(\mathcal{P})$ (namely the set $\mathcal{L}(A, T)$) to the vertices in $\Gamma(\mathcal{Q})$. Therefore when we extend ϕ to edges and paths the definition will only extend to those edges and paths in $\Gamma(\mathcal{P})$ that begin in $\mathcal{L}(A, T)$.

Now we look to extend this idea as we think of the derivation graphs for $\Gamma(\mathcal{P})$ and $\Gamma(\mathcal{Q})$. Here the words are vertices and we have edges between words. Recall that the edges represent a single application of a relation to a string of letters. The free monoid A^* acts on $\Gamma(\mathcal{P})$ and the free monoid B^* acts on $\Gamma(\mathcal{Q})$, extending any edge on the left and or the right by multiplication (which in our case is defined as concatenation). This means that each relation from the presentation is one of the edges but can itself be embedded in a much longer word, and with other relations, to make up many more edges. A path within the graph represents a sequence of single applications of relations to a word.

Definition 4.6.2. Let $\mathbb{E} \in \Gamma(\mathcal{P})$ be an arbitrary edge with $\iota\mathbb{E} \in \mathcal{L}(A,T)$ and $\mathbb{E} = (w_1, r, \epsilon, w_2)$ with $w_1, w_2 \in A^*$ and $r \in R$. Let $\iota\mathbb{E} \equiv w_1 r_\epsilon w_2 \equiv w_1' w_1'' r_\epsilon w_2' w_2''$ where $w_1', w_1'', w_2', w_2'' \in A^*$ and satisfying the following properties:

(a) w_1' is the longest prefix of w_1 in $\mathcal{L}(A,T)$;

(b) $w_1 r_\epsilon w_2'$ is the shortest prefix of $w_1 r_\epsilon w_2$ which is in $\mathcal{L}(A,T)$ and has $w_1 r_\epsilon$ as a prefix.

Then we define

$$\mathbb{E}\phi = (w_1'\phi, (w_1'' r_{+1} w_2')\phi = (w_1'' r_{-1} w_2')\phi, \epsilon, w_2''\phi).$$

<u>Remark</u>:

When we recall the definition of ϕ we can see that it is defined on words that belong to $\mathcal{L}(A,T)$. Here we give an explanation of how ϕ is applied to the edge $\mathbb{E}$ in Definition 4.6.2 above, in particular we show that w_1', $w_1'' r_{+1} w_2'$, $w_1'' r_{-1} w_2'$ and w_2'' all belong to $\mathcal{L}(A,T)$. The explanation below is for positive edges, so we assume $\epsilon = +1$, similar arguments apply for negative edges where $\epsilon = -1$.

(1) w_1' belongs to $\mathcal{L}(A,T)$ by part (a) in the definition.

(2) $w_1'' r_{+1} w_2'$ belongs to $\mathcal{L}(A,T)$ since $w_1' \in \mathcal{L}(A,T)$ and $w_1 r_{+1} w_2' \in \mathcal{L}(A,T)$ by part (b), so since T is a left unitary subsemigroup and $w_1 r_{+1} w_2' \equiv w_1' w_1'' r_{+1} w_2'$ it follows that $w_1'' r_{+1} w_2' \in \mathcal{L}(A,T)$.

(3) $w_1'' r_{-1} w_2'$ belongs to $\mathcal{L}(A,T)$ since $w_1'' r_{+1} w_2' \in \mathcal{L}(A,T)$ by part (2) and $w_1'' r_{-1} w_2' = w_1'' r_{+1} w_2'$ in T.

(4) w_2'' belongs to $\mathcal{L}(A,T)$ as follows. By assumption $\iota\mathbb{E} \in \mathcal{L}(A,T)$ and $\iota\mathbb{E} \equiv w_1 r_{+1} w_2$. Since $w_1 r_{+1} w_2' \in \mathcal{L}(A,T)$ and $w_1 r_{+1} w_2' w_2'' \equiv w_1 r_{+1} w_2$ is in $\mathcal{L}(A,T)$ and T is a left unitary semigroup, it follows that w_2'' belongs to $\mathcal{L}(A,T)$ as required.

Lemma 4.6.3. *Let $\mathbb{E} \in \Gamma(\mathcal{P})$ be an arbitrary edge with $\iota\mathbb{E} \in \mathcal{L}(A,T)$ and $\mathbb{E} = (w_1, r, \epsilon, w_2)$. Decompose $w_1 \equiv w_1'w_1''$ and $w_2 \equiv w_2'w_2''$ as in Definition 4.6.2. Then*

(i) $(\iota\mathbb{E})\phi, (\tau\mathbb{E})\phi \in B^+$, $(\iota\mathbb{E})\phi \equiv \iota(\mathbb{E}\phi)$ *and* $\tau(\mathbb{E}\phi) \equiv (\tau\mathbb{E})\phi$;

(ii) $(w_1''r_{+1}w_2')\phi = (w_1''r_{-1}w_2')\phi \in U$ *in the presentation* $\langle B \mid U\rangle$;

(iii) $\mathbb{E}\phi$ *is an edge in the derivation graph* $\Gamma(\mathcal{Q})$;

(iv) $\mathbb{E}\phi = w_1'\phi \cdot \mathbb{E}_1\phi \cdot w_2''\phi$ *where* $\mathbb{E}_1 = (1, (w_1''r_{+1}w_2') = (w_1''r_{-1}w_2'), \epsilon, 1) \in \Gamma(\mathcal{P})$ *and* $\iota\mathbb{E}_1 \in \mathcal{L}(A,T)$;

(v) $\mathbb{E}^{-1}\phi = (\mathbb{E}\phi)^{-1}$ *in* $\Gamma(\mathcal{Q})$.

Proof. We shall prove the result for positive edges, so we assume $\epsilon = +1$. The proof for negative edges then follows from this together with the definitions. The proof is given for each part as follows:

(i) By Definition 4.6.2 and ϕ we have $\iota(\mathbb{E}\phi) \equiv (w_1'\phi)\big((w_1''r_{+1}w_2')\phi\big)(w_2''\phi)$ and $(\iota\mathbb{E})\phi \equiv (w_1r_{+1}w_2)\phi$. If we apply ϕ as in Definition 4.6.2 and decompose $(w_1r_{+1}w_2)$, then by Lemma 4.5.5 we have $w_1r_{+1}w_2 \equiv w_1'w_1''r_{+1}w_2'w_2''$ where w_1', $w_1''r_{+1}w_2'$ and w_2'' are all in $\mathcal{L}(A,T)$. Hence, $(\iota\mathbb{E})\phi \equiv (w_1r_{+1}w_2)\phi \equiv (w_1'w_1''r_{+1}w_2'w_2'')\phi \equiv (w_1'\phi)\big((w_1''r_{+1}w_2')\phi\big)(w_2''\phi) \equiv \iota(\mathbb{E}\phi)$. Therefore, $\iota(\mathbb{E}\phi) \equiv (\iota\mathbb{E})\phi \in B^+$. Similarly for $\tau(\mathbb{E}\phi)$ and $(\tau\mathbb{E})\phi$.

(ii) We claim that, when $w_1''r_{+1}w_2'$ is decomposed into principal factors (see Lemma 4.5.5 and the proof of Theorem 4.5.2) each of these factors must contain at least one letter from r_{+1}, thus $|(w_1''r_{+1}w_2')\phi| \leq |r_{+1}|$.

To prove the claim, consider decomposing $w_1''r_{+1}w_2'$ into principal factors as in Lemma 4.5.5 and the proof of Theorem 4.5.2. We obtain $w_1''r_{+1}w_2' \equiv \alpha_1\alpha_2\ldots\alpha_k$ where each α_i belongs to $\mathcal{SWB}_r(A,T)$. Now since w_1' is the longest prefix of w_1 in $\mathcal{L}(A,T)$ it follows that w_1'' cannot have a non-empty prefix in $\mathcal{L}(A,T)$, hence α_1 must contain w_1'' as a proper prefix. Also, if α_k

were a suffix of w_2' then $w_1'\alpha_1\alpha_2\dots\alpha_{k-1}$ would be a prefix of $w_1r_{+1}w_2$ which is in $\mathcal{L}(A,T)$ and contains w_1r_{+1} as a prefix. But this would contradict the fact that $w_1r_{+1}w_2'$ is the shortest such prefix of $w_1r_{+1}w_2$. Hence α_k must contain w_2' as a proper suffix. It follows that each α_i must contain at least one letter of r_{+1}, proving the claim.

A similar argument applies to $(w_1''r_{-1}w_2')\phi$ and we have $|(w_1''r_{-1}w_2')\phi| \leq |r_{-1}|$. Therefore $|(w_1''r_{+1}w_2'w_1''r_{-1}w_2')\phi| \leq |r_{+1}r_{-1}|$ where $(r_{+1},r_{-1}) \in R$ and so by Theorem 4.5.6 $(w_1''r_{+1}w_2')\phi = (w_1''r_{-1}w_2')\phi \in U$.

(iii) By Definition 4.6.2, $\mathbb{E}\phi = (w_1'\phi, (w_1''r_{+1}w_2')\phi = (w_1''r_{-1}w_2')\phi, +1, w_2''\phi)$. By the definition of ϕ both $w_1'\phi$ and $w_2''\phi$ are in B^*. Note that both w_1' and w_2'' could be the empty word. By part (ii) $(w_1''r_{+1}w_2')\phi = (w_1''r_{-1}w_2')\phi \in U$. Thus $\mathbb{E}\phi$ is an edge in the graph $\Gamma(\mathcal{Q})$.

(iv) Follows from the definition of an edge in Section 4.4 and part (ii).

(v) We prove the result for the case that $\mathbb{E}$ is a positive edge, the other case being similar. So we have $\mathbb{E} = (w_1, r, +1, w_2)$ and $\mathbb{E}^{-1} = (w_1, r, -1, w_2)$ where $\iota\mathbb{E} \equiv \tau\mathbb{E}^{-1} \equiv w_1r_{+1}w_2$ and $\tau\mathbb{E} \equiv \iota\mathbb{E}^{-1} \equiv w_1r_{-1}w_2$. By assumption $w_1r_{+1}w_2 \in \mathcal{L}(A,T)$, which is equivalent to $w_1r_{-1}w_2 \in \mathcal{L}(A,T)$ since $w_1r_{+1}w_2 = w_1r_{-1}w_2$ in the semigroup. We want to prove that $(\mathbb{E}\phi)^{-1} = \mathbb{E}^{-1}\phi$.

By definition $\mathbb{E}\phi = \big(w_1'\phi, (w_1''r_{+1}w_2')\phi = (w_1''r_{-1}w_2')\phi, +1, w_2''\phi\big)$ where

(a) w_1' is the longest prefix of w_1 in $\mathcal{L}(A,T)$;

(b) $w_1r_{+1}w_2'$ is the shortest prefix of $w_1r_{+1}w_2$ which is in $\mathcal{L}(A,T)$ and has w_1r_{+1} as a prefix.

Hence it follows that $(\mathbb{E}\phi)^{-1} = \big(w_1'\phi, (w_1''r_{+1}w_2')\phi = (w_1''r_{-1}w_2')\phi, -1, w_2''\phi\big)$.

Now consider $\mathbb{E}^{-1}\phi$. Since $\iota\mathbb{E}^{-1} \equiv w_1r_{-1}w_2$ it follows from Definition 4.6.2 that $\mathbb{E}^{-1}\phi = \big(u_1'\phi, (u_1''r_{+1}u_2')\phi = (u_1''r_{-1}u_2')\phi, -1, u_2''\phi\big)$ where

(a) u_1' is the longest prefix of w_1 in $\mathcal{L}(A,T)$;

(b) $w_1r_{-1}u_2'$ is the shortest prefix of $w_1r_{-1}w_2$ which is in $\mathcal{L}(A,T)$ and has w_1r_{-1} as a prefix.

It is immediate from the definition that $u_1' \equiv w_1'$ and hence $u_1'' \equiv w_1''$. Furthermore, since $w_1r_{+1} = w_1r_{-1}$ in the semigroup S it follows that for any prefix v of w_2 we have $w_1r_{+1}v = w_1r_{-1}v$ and hence $w_1r_{+1}v \in \mathcal{L}(A,T) \Leftrightarrow w_1r_{-1}v \in \mathcal{L}(A,T)$. It follows from this that $w_2' \equiv u_2'$ and hence $w_2'' \equiv u_2''$. This completes the proof that $(\mathbb{E}\phi)^{-1} = \mathbb{E}^{-1}\phi$.

□

Definition 4.6.4. Let $\mathbb{P}$ be an arbitrary path in $\Gamma(\mathcal{P})$ with $\mathbb{P} = \mathbb{E}_1 \ldots \mathbb{E}_k$, $k \geq 1$ and $\iota\mathbb{P} \in \mathcal{L}(A,T)$. Then we define

$$\mathbb{P}\phi = (\mathbb{E}_1\phi)\ldots(\mathbb{E}_k\phi).$$

We claim that if $\mathbb{P}$ is a path in $\Gamma(\mathcal{P})$ with $\iota\mathbb{P}$ in $\mathcal{L}(A,T)$ then $\mathbb{P}\phi$ is a path in $\Gamma(\mathcal{Q})$. We shall prove this in the next lemma. In the following two lemmas we also record some other important facts about the behaviour of the function ϕ when applied to such paths.

Lemma 4.6.5. *Let $\mathbb{P}$ be an arbitrary path in $\Gamma(\mathcal{P})$ with $\mathbb{P} = \mathbb{E}_1 \ldots \mathbb{E}_k$, $k \geq 1$ and $\iota\mathbb{P} \in \mathcal{L}(A,T)$. Then*

(i) $\tau(\mathbb{E}_i\phi) \equiv \iota(\mathbb{E}_{i+1}\phi)$ *for all* $i \in \{1,\ldots,k-1\}$*;*

(ii) $(\iota\mathbb{P})\phi \equiv \iota(\mathbb{P}\phi)$ *and* $(\tau\mathbb{P})\phi \equiv \tau(\mathbb{P}\phi)$*;*

(iii) $\mathbb{P}\phi$ *is a path in* $\Gamma(\mathcal{Q})$*;*

(iv) $\mathbb{P}^{-1}\phi = (\mathbb{E}_k^{-1}\ldots\mathbb{E}_1^{-1})\phi = (\mathbb{E}_k)^{-1}\phi\ldots(\mathbb{E}_1)^{-1}\phi = (\mathbb{P}\phi)^{-1}$.

Proof. The proof is given for each part as follows:

(i) In the path $\mathbb{P} \in \Gamma(\mathcal{P})$ we have, by the definition of paths in Section 4.4, $\tau\mathbb{E}_i \equiv \iota\mathbb{E}_{i+1}$ for all $i \in \{1,\ldots,k-1\}$. It then immediately follows that

$(\tau\mathbb{E}_i)\phi \equiv (\iota\mathbb{E}_{i+1})\phi$ in $\Gamma(\mathcal{Q})$. From Lemma 4.6.3 part (i) we then have $\tau(\mathbb{E}_i\phi) \equiv (\tau\mathbb{E}_i)\phi \equiv (\iota\mathbb{E}_{i+1})\phi \equiv \iota(\mathbb{E}_{i+1}\phi)$ as required.

(ii) We have $\iota\mathbb{P} \equiv \iota\mathbb{E}_1$ by the definition of paths in Section 4.4 which gives $(\iota\mathbb{P})\phi \equiv (\iota\mathbb{E}_1)\phi$. Also $(\iota\mathbb{E}_1)\phi \equiv \iota(\mathbb{E}_1\phi)$ by Lemma 4.6.3 part (i). So now $(\iota\mathbb{P})\phi \equiv \iota(\mathbb{E}_1\phi) \equiv \iota(\mathbb{P}\phi)$, the last equality by the definition of ϕ and paths. Similarly $(\tau\mathbb{P})\phi \equiv (\tau\mathbb{E}_k)\phi \equiv \tau(\mathbb{E}_k\phi) \equiv \tau(\mathbb{P}\phi)$ in $\Gamma(\mathcal{Q})$.

(iii) By Lemma 4.6.3 for every edge $\mathbb{E}_i$ in the path $\mathbb{P}$ we have $\mathbb{E}_i\phi$ is an edge in $\Gamma(\mathcal{Q})$. Together with Lemma 4.6.5 part (i) it then follows that $\mathbb{P}\phi$ is a path in $\Gamma(\mathcal{Q})$.

(iv) We have

$$\begin{aligned}
\mathbb{P}^{-1}\phi &= (\mathbb{E}_1 \dots \mathbb{E}_k)^{-1}\phi \\
&= (\mathbb{E}_k^{-1} \dots \mathbb{E}_1^{-1})\phi && [\text{ by the definition of paths }] \\
&= (\mathbb{E}_k^{-1}\phi) \dots (\mathbb{E}_1^{-1}\phi) && [\text{ by the definition of } \phi\ ] \\
&= (\mathbb{E}_k\phi)^{-1} \dots (\mathbb{E}_1\phi)^{-1} && [\text{ by Lemma 4.6.3 (v) }] \\
&= (\mathbb{E}_1\phi \dots \mathbb{E}_k\phi)^{-1} && [\text{ by the definition of paths }] \\
&= (\mathbb{P}\phi)^{-1}.
\end{aligned}$$

□

Lemma 4.6.6. *Let $\mathbb{P}$ be a path in $\Gamma(\mathcal{P})$ with $\iota\mathbb{P} \in \mathcal{L}(A,T)$ and let $w_1, w_2 \in \mathcal{L}(A,T)$. Then*

$$(w_1 \cdot \mathbb{P} \cdot w_2)\phi = w_1\phi \cdot \mathbb{P}\phi \cdot w_2\phi.$$

Proof. Write $\mathbb{P} = \mathbb{E}_1\mathbb{E}_2 \dots \mathbb{E}_k$ where $\mathbb{E}_i \in E(\Gamma(\mathcal{P}))$ for $1 \leq i \leq k$. Since $\iota\mathbb{P} \in \mathcal{L}(A,T)$ it follows that $\iota\mathbb{E}_i \in \mathcal{L}(A,T)$ for all $1 \leq i \leq k$. Then by definition of ϕ on edges, and since $w_1, w_2 \in \mathcal{L}(A,T)$ it follows that

$$(w_1 \cdot \mathbb{E}_i \cdot w_2)\phi = w_1\phi \cdot (\mathbb{E}_i\phi) \cdot w_2\phi$$

for all $i \in \{1, 2, \ldots, k\}$. Hence

$$\begin{aligned}(w_1 \cdot \mathbb{P} \cdot w_2)\phi &= [(w_1 \cdot \mathbb{E}_1 \cdot w_2) \circ (w_1 \cdot \mathbb{E}_2 \cdot w_2) \circ \ldots \circ (w_1 \cdot \mathbb{E}_k \cdot w_2)]\phi \\ &= [(w_1 \cdot \mathbb{E}_1 \cdot w_2)\phi] \circ \ldots \circ [(w_1 \cdot \mathbb{E}_k \cdot w_2)\phi] \\ &= [w_1\phi \cdot (\mathbb{E}_1\phi) \cdot w_2\phi] \circ \ldots \circ [w_1\phi \cdot (\mathbb{E}_k\phi) \cdot w_2\phi] \\ &= w_1\phi \cdot (\mathbb{P}\phi) \cdot w_2\phi,\end{aligned}$$

as required. □

Recall that we use $P(\Gamma(\mathcal{P}))$ to denote the set of all paths in $\Gamma(\mathcal{P})$, and similarly $P(\Gamma(\mathcal{Q}))$ for the set of paths in $\Gamma(\mathcal{Q})$. Also recall that if $w \in A^+$ then 1_w denotes the empty path at w in $\Gamma(\mathcal{P})$, and similarly for 1_u with $u \in B^+$. Recall that $\mathcal{P} = \langle A \mid R \rangle$ and $\mathcal{Q} = \langle B \mid U \rangle$. The next definition extends ψ to a mapping which sends paths in $\Gamma(\mathcal{Q})$ to paths in $\Gamma(\mathcal{P})$.

Definition 4.6.7. We extend $\psi : B^+ \to A^+$ to a mapping, also denoted ψ, where

$$\psi : P(\Gamma(\mathcal{Q})) \to P(\Gamma(\mathcal{P})).$$

To do this, firstly for any word $w \in B^+$ we define $1_w\psi = 1_{w\psi}$. Next, for each $u \in U$ let $\mathbb{E}_u = \{1, u, +1, 1)$, which is an edge in $\Gamma(\mathcal{Q})$. Then for every $u \in U$ let $\mathbb{E}_u\psi$ be a fixed path in $\Gamma(\mathcal{P})$ from $u_{+1}\psi$ to $u_{-1}\psi$. Such a path exists since $u_{+1}\psi$ and $u_{-1}\psi$ represent the same element of S. Then for an arbitrary edge $\mathbb{E} = (w_1, u, \epsilon, w_2)$ of $\Gamma(\mathcal{Q})$ we define

$$\mathbb{E}\psi = (w_1\psi) \cdot \big((\mathbb{E}_u)\psi\big)^{\epsilon} \cdot (w_2\psi).$$

Then for any path $\mathbb{P} = \mathbb{E}_1\mathbb{E}_2 \ldots \mathbb{E}_k$ in $\Gamma(\mathcal{Q})$ we define

$$\mathbb{P}\psi = (\mathbb{E}_1\psi)(\mathbb{E}_2\psi) \ldots (\mathbb{E}_k\psi).$$

The following lemma is an immediate consequence of Definition 4.6.7, so we omit the proof.

Lemma 4.6.8. *Let $\mathbb{P} \in P(\Gamma(\mathcal{Q}))$ be a path in $\Gamma(\mathcal{Q})$. Then*

(i) $\mathbb{P}\psi$ *is a path in* $\Gamma(\mathcal{P})$ *with* $\iota(\mathbb{P}\psi) \equiv (\iota\mathbb{P})\psi$ *and* $\tau(\mathbb{P}\psi) \equiv (\tau\mathbb{P})\psi$;

(ii) $(\mathbb{P}^{-1})\psi = (\mathbb{P}\psi)^{-1}$.

The key lemma we need to prove next is the following.

Lemma 4.6.9. *For every edge* $\mathbb{E} = (w_1, u, \epsilon, w_2) \in E(\Gamma(\mathcal{Q}))$ *we have*

$$\mathbb{E}\psi\phi = w_1 \cdot (\mathbb{E}_u\psi\phi)^{\epsilon} \cdot w_2 \text{ where } \mathbb{E}_u = (1, u, +1, 1).$$

Proof. We prove the result in the case $\epsilon = +1$, the other case being similar.

We have

$\mathbb{E}\psi\phi$

$= [(w_1\psi) \cdot (\mathbb{E}_u\psi) \cdot (w_2\psi)]\phi$ (by definition of $\mathbb{E}\psi$)

$= (w_1\psi\phi) \cdot (\mathbb{E}_u\psi\phi) \cdot (w_2\psi\phi)$

(by Lemma 4.6.6, since $w_1\psi, w_2\psi \in \mathcal{L}(A,T)$ and $\iota(\mathbb{E}_u\psi) \equiv (\iota\mathbb{E}_u)\psi \in \mathcal{L}(A,T)$)

$= w_1 \cdot (\mathbb{E}_u\psi\phi) \cdot w_2$ (since $w_1\psi\phi \equiv w_1$ and $w_2\psi\phi \equiv w_2$ by Lemma 4.6.1). □

4.6.4 Create an infinite homotopy base for the derivation graph $\Gamma(\mathcal{Q})$

Next we define an infinite homotopy base for the derivation graph $\Gamma(\mathcal{Q})$ of the presentation $\mathcal{Q} = \langle B \mid U \rangle$.

Another thing we need in order to write down an infinite homotopy base for $\Gamma(\mathcal{Q})$ is, for each $w \in B^+$, to choose and fix a path Λ_w in $\Gamma(\mathcal{Q})$ with initial vertex w and terminal vertex $w\psi\phi$. Since by Lemma 4.6.1 we in fact have $w \equiv w\psi\phi$ for all $w \in B^+$ we can and will simply define Λ_w to be the empty path 1_w at w for all $w \in B^+$. So for the remainder of this section for all $w \in B^+$ we use Λ_w to denote the empty path in $\Gamma(\mathcal{Q})$ at w. While this might seem superfluous (nothing would change below if all occurrences of Λ_w were deleted) it will make it easier for us to make reference to general results from the literature about homotopy bases, where Λ notation is used.

Since the semigroup S has FDT, there exists a finite homotopy base for $\Gamma(\mathcal{P})$. Let X be a finite homotopy base for the presentation $\mathcal{P} = \langle A \mid R\rangle$ which defines the semigroup S, where X is a set of closed paths. Let Z denote the following infinite set of closed paths for the graph $\Gamma(\mathcal{Q})$:

(4) $\mathbb{E}\Lambda_{\tau\mathbb{E}}\big((\mathbb{E}\psi)\phi\big)^{-1}\Lambda_{\iota\mathbb{E}}^{-1}$, for $\mathbb{E}$ in $\Gamma(\mathcal{Q})$;

(5) $([\mathbb{E}_1, \mathbb{E}_2])\phi$, for $\mathbb{E}_1, \mathbb{E}_2$ in $\Gamma(\mathcal{P})$ such that $\iota\mathbb{E}_1\iota\mathbb{E}_2 \in \mathcal{L}(A, T)$, where $[\mathbb{E}_1, \mathbb{E}_2]$ is the path $(\mathbb{E}_1 \cdot \iota\mathbb{E}_2)(\tau\mathbb{E}_1 \cdot \mathbb{E}_2)(\mathbb{E}_1 \cdot \tau\mathbb{E}_2)^{-1}(\iota\mathbb{E}_1 \cdot \mathbb{E}_2)^{-1}$;

(6) $(w_1 \cdot \mathbb{P} \cdot w_2)\phi$, for $\mathbb{P} \in X$ and $w_1, w_2 \in A^*$ such that $w_1(\iota\mathbb{P})w_2 \in \mathcal{L}(A, T)$.

Lemma 4.6.10. *Let Z be the set of closed paths as defined above. Then Z is an infinite homotopy base for the presentation $\langle B \mid U\rangle$ which defines the subsemigroup T.*

Proof. Using a standard argument (see [19, Lemma 9]) it may be shown that the set of closed paths Z is an infinite homotopy base for $\Gamma(\mathcal{Q})$. This follows from a more general result in [39]. $\square$

The key now is to define a finite set of closed paths for which the closed paths of types (4), (5) and (6) are a consequence. Then we will have proved that $\Gamma(\mathcal{Q})$ has a finite homotopy base and thus has FDT.

4.6.5 Create a finite homotopy base for the derivation graph $\Gamma(\mathcal{Q})$

We will consider each of the path types (4), (5) and (6) in turn.

Type (4) paths.

We consider the infinite set of closed paths of type (4) where for any $\mathbb{E} \in E(\Gamma(\mathcal{Q}))$ we take the closed path $\mathbb{E}\Lambda_{\tau\mathbb{E}}((\mathbb{E}\psi)\phi)^{-1}\Lambda_{\iota\mathbb{E}}^{-1}$. The following Diagram 4.6.1 illustrates a general case for a type (4) path. Note that in our case we have chosen Λ_w to be the empty path 1_w for all $w \in B^+$.

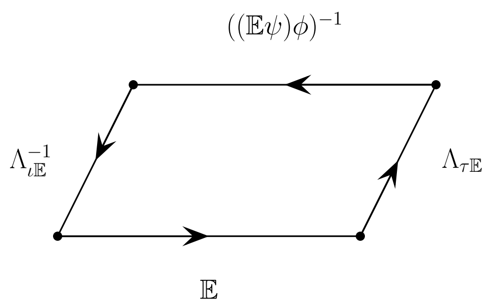


Figure 4.6.1: A general diagram for Type (4) paths.

Lemma 4.6.11. *Let*

$$W = \{\ \mathbb{E}_u(\mathbb{E}_u\psi\phi)^{-1} : u \in U, \mathbb{E}_u = (1, u, +1, 1)\ \}.$$

Then for every edge $\mathbb{E}$ *in* $\Gamma(\mathcal{Q})$ *we have*

$$\mathbb{E}\ \Lambda_{\tau\mathbb{E}}\ (\mathbb{E}\psi\phi)^{-1}\ \Lambda_{\iota\mathbb{E}}^{-1}\ \sim_W\ 1_{\iota\mathbb{E}}.$$

Proof. Since by definition $\Lambda_{\tau\mathbb{E}}$ and $\Lambda_{\iota\mathbb{E}}$ are both empty paths, this is equivalent to proving that $\mathbb{E} \sim_W \mathbb{E}\psi\phi$. Write $\mathbb{E} = (w_1, u, \epsilon, w_2)$. Then by Lemma 4.6.9 we have:

$$\mathbb{E}\psi\phi = w_1 \cdot (\mathbb{E}_u\psi\phi)^{\epsilon} \cdot w_2 \text{ where } \mathbb{E}_u = (1, u, +1, 1).$$

By definition we have $\mathbb{E} = w_1 \cdot \mathbb{E}_u^{\epsilon} \cdot w_2$.
By definition of W we have

$$\mathbb{E}_u \sim_W \mathbb{E}_u\psi\phi \text{ and hence also } \mathbb{E}_u^{-1} \sim_W (\mathbb{E}_u\psi\phi)^{-1}.$$

Therefore

$$\mathbb{E} = w_1 \cdot \mathbb{E}_u^{\epsilon} \cdot w_2 \sim_W w_1 \cdot (\mathbb{E}_u\psi\phi)^{\epsilon} \cdot w_2 = \mathbb{E}\psi\phi,$$

and this completes the proof of the lemma. $\square$

Remark: Since the set of all relations U in the presentation $\mathcal{Q} = \langle B \mid U \rangle$ is finite, it follows that W defined in Lemma 4.6.11 is a finite set of closed paths in $\Gamma(\mathcal{Q})$.

Type (5) paths.

We consider the infinite set of type (5) paths $([\mathbb{E}_1, \mathbb{E}_2])\phi$, for $\mathbb{E}_1, \mathbb{E}_2$ in $\Gamma(\mathcal{P})$ such that $\iota\mathbb{E}_1\iota\mathbb{E}_2 \in \mathcal{L}(A, T)$, where

$[\mathbb{E}_1, \mathbb{E}_2]$ is the closed path $(\mathbb{E}_1 \cdot \iota\mathbb{E}_2)(\tau\mathbb{E}_1 \cdot \mathbb{E}_2)(\mathbb{E}_1 \cdot \tau\mathbb{E}_2)^{-1}(\iota\mathbb{E}_1 \cdot \mathbb{E}_2)^{-1}$.

First we make some useful definitions.

Definition 4.6.12. Let $\mathbb{E}_1 = (w_1, r, \epsilon, w_2)$ and $\mathbb{E}_2 = (u_1, s, \delta, u_2)$ where $\epsilon, \delta \in \{-1, +1\}$ such that $\iota\mathbb{E}_1\iota\mathbb{E}_2 \in \mathcal{L}(A, T)$ with $w_1, w_2, u_1, u_2 \in A^*$ and $r, s \in R$. We say that *a word* $v \in A^+$ *splits* $[\mathbb{E}_1, \mathbb{E}_2]$ if

(a) $v \in \mathcal{L}(A, T)$ and

(b) v contains $w_1 r_+$ as a prefix and v is a prefix of $w_1 r_+ w_2 u_1$.

We say that $[\mathbb{E}_1, \mathbb{E}_2]$ *splits* if there is such a word v which splits it.

Now we can look at the infinite set of paths $([\mathbb{E}_1, \mathbb{E}_2])\phi$ and there are two cases to consider,

Case 1: $[\mathbb{E}_1, \mathbb{E}_2]$ splits (this is illustrated in Figure 4.6.2 below) and

Case 2: $[\mathbb{E}_1, \mathbb{E}_2]$ does not split (this is illustrated in Figure 4.6.3 below).

First we consider Case 1.

Case 1: $[\mathbb{E}_1, \mathbb{E}_2]$ splits.

In the following Figure 4.6.2 we illustrate Case 1 with the split. In particular, the split is within the word w_2 which means that $w_3 \equiv w_2'$ and $u_3 \equiv w_2'' u_1$ in the proof. The split could equally occur within the word u_1 or immediately at the start or end of the words w_2 and u_1. The notation used in the diagram matches the notation used in the proof of Lemma 4.6.13 below.

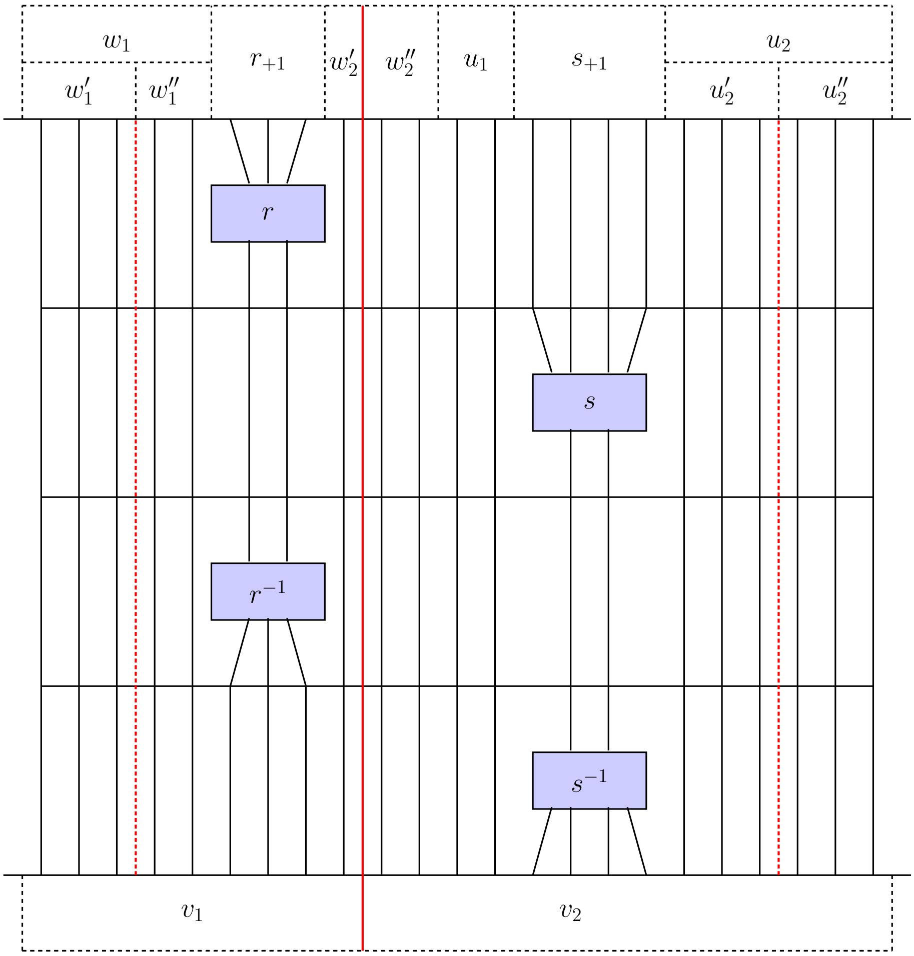


Figure 4.6.2: Type (5) paths Case 1 with split (shown by a solid red line).

In the following Lemma 4.6.13, and throughout this section, we use the notation $\sim$ to denote $\sim_{\varnothing}$ i.e. when we write $\mathbb{P} \sim \mathbb{Q}$ it means that $\mathbb{P}$ and $\mathbb{Q}$ are $\sim$ -related modulo the empty set.

Lemma 4.6.13. *Let $\mathbb{E}_1, \mathbb{E}_2$ be edges in $\Gamma(\mathcal{P})$ with $w \equiv \iota\mathbb{E}_1\iota\mathbb{E}_2 \in \mathcal{L}(A,T)$. If $[\mathbb{E}_1, \mathbb{E}_2]$ splits then $[\mathbb{E}_1, \mathbb{E}_2]\phi \sim 1_{w\phi}$ in $\Gamma(\mathcal{Q})$.*

Proof. We prove the result in the case that $\mathbb{E}_1$ and $\mathbb{E}_2$ are positive edges. The other cases can be dealt with using the same argument. Let $\mathbb{E}_1 = (w_1, r, +1, w_2)$ and $\mathbb{E}_2 = (u_1, s, +1, u_2)$ with $w_1, w_2, u_1, u_2 \in A^*$ and $r, s \in R$. Let the word v_1 split $[\mathbb{E}_1, \mathbb{E}_2]$ such that $w \equiv v_1v_2$ where $v_1 \equiv w_1r_+w_3$ and $v_2 \equiv u_3s_+u_2$ with $w_3, u_3 \in A^*$. Note that $v_1, v_2 \in \mathcal{L}(A,T)$.

Next we can rewrite $[\mathbb{E}_1, \mathbb{E}_2]$ to reflect the split. Let $\mathbb{E}_1' = (w_1, r, +1, w_3)$ and $\mathbb{E}_2' = (u_3, s, +1, u_2)$ then $[\mathbb{E}_1, \mathbb{E}_2] = [\mathbb{E}_1', \mathbb{E}_2']$ and hence $[\mathbb{E}_1, \mathbb{E}_2]\phi = [\mathbb{E}_1', \mathbb{E}_2']\phi$. It is important to note that $\iota\mathbb{E}_1'$, $\tau\mathbb{E}_1'$, $\iota\mathbb{E}_2'$ and $\tau\mathbb{E}_2'$ all belong to $\mathcal{L}(A, T)$. Since $\iota\mathbb{E}_1'$, $\tau\mathbb{E}_1'$, $\iota\mathbb{E}_2'$ and $\tau\mathbb{E}_2'$ all belong to $\mathcal{L}(A, T)$, it follows that

$[\mathbb{E}_1', \mathbb{E}_2']\phi$
$= [\ (\mathbb{E}_1' \cdot \iota\mathbb{E}_2')(\tau\mathbb{E}_1' \cdot \mathbb{E}_2')(\mathbb{E}_1' \cdot \tau\mathbb{E}_2')^{-1}(\iota\mathbb{E}_1' \cdot \mathbb{E}_2')^{-1}\ ]\phi$
(by definition of the path $[\mathbb{E}_1', \mathbb{E}_2']$),

$= [(\mathbb{E}_1' \cdot \iota\mathbb{E}_2')\phi]\ [(\tau\mathbb{E}_1' \cdot \mathbb{E}_2')\phi]\ [(\mathbb{E}_1' \cdot \tau\mathbb{E}_2')^{-1}\phi]\ [(\iota\mathbb{E}_1' \cdot \mathbb{E}_2')^{-1}\phi]$
(by Definition 4.6.4),

$= [(\mathbb{E}_1'\phi) \cdot (\iota\mathbb{E}_2'\phi)]\ [(\tau\mathbb{E}_1'\phi) \cdot (\mathbb{E}_2'\phi)]\ [(\mathbb{E}_1'\phi)^{-1} \cdot (\tau\mathbb{E}_2'\phi)]\ [(\iota\mathbb{E}_1'\phi) \cdot (\mathbb{E}_2'\phi)^{-1}]$
(by Lemma 4.6.6),

$= [(\mathbb{E}_1'\phi) \cdot \iota(\mathbb{E}_2'\phi)]\ [\tau(\mathbb{E}_1'\phi) \cdot (\mathbb{E}_2'\phi)]\ [(\mathbb{E}_1'\phi)^{-1} \cdot \tau(\mathbb{E}_2'\phi)]\ [\iota(\mathbb{E}_1'\phi) \cdot (\mathbb{E}_2'\phi)^{-1}]$
(by Lemmas 4.6.5 (ii) and 4.6.3 (i)),

$= [\mathbb{E}_1'\phi, \mathbb{E}_2'\phi] \sim 1_{w\phi}$
(since $\mathbb{E}_1'\phi \in E(\Gamma(\mathcal{Q}))$ and $\mathbb{E}_2'\phi \in E(\Gamma(\mathcal{Q}))$).

Hence we have $[\mathbb{E}_1, \mathbb{E}_2]\phi = [\mathbb{E}_1', \mathbb{E}_2']\phi = [\mathbb{E}_1'\phi, \mathbb{E}_2'\phi] \sim 1_{w\phi}$.
This completes the proof. □

Case 2: $[\mathbb{E}_1, \mathbb{E}_2]$ does not split.

In the following Figure 4.6.3 we illustrate Case 2. In particular, we have illustrated the case where we have a decomposition of the word $w_1'' r_{+1} w_2 u_1 s_{+1} u_2'$ into the maximum number of possible principal factors (which in some instances are individual letters), each of which belong to $\mathcal{L}(A, T)$. See Lemma 4.5.5 and the proof of Theorem 4.5.2 for details of decomposition into principal factors. So we have $w_1'' r_{+1} w_2 u_1 s_{+1} u_2' \equiv \alpha_1\alpha_2\alpha_3\alpha_4\alpha_5\alpha_6$ where α_1, α_2, α_3, α_4, α_5 and α_6 all belong to $\mathcal{L}(A, T)$. Note that the words w_1'' and u_2' contain no prefix that belong to $\mathcal{L}(A, T)$ by definition of the edges $\mathbb{E}_1$ and $\mathbb{E}_2$. The notation used in the diagram matches the notation used in the proofs of Lemma 4.6.14 and Lemma 4.6.15 below.

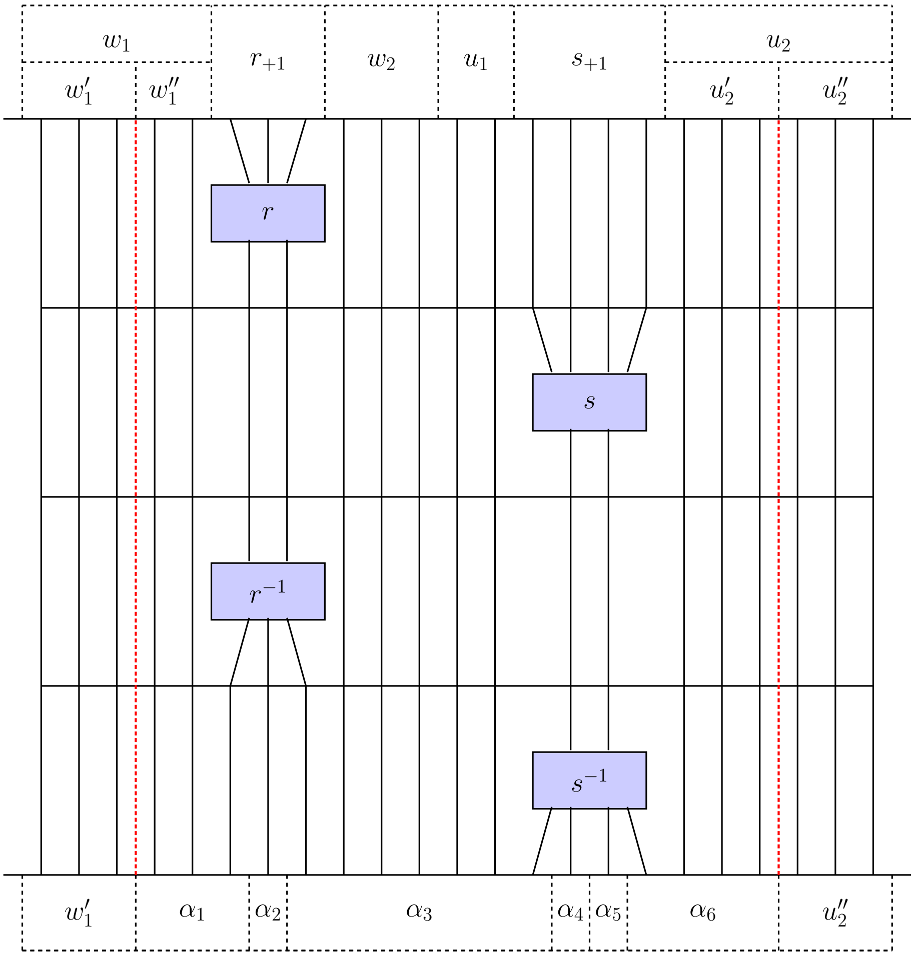


Figure 4.6.3: Type (5) paths Case 2 no split

Lemma 4.6.14. *Let* $\mathbb{E}_1, \mathbb{E}_2$ *be edges in* $\Gamma(\mathcal{P})$ *with* $w = \iota\mathbb{E}_1\iota\mathbb{E}_2 \in \mathcal{L}(A,T)$. *If* $[\mathbb{E}_1, \mathbb{E}_2]$ *does not split then there are edges* $\mathbb{F}_1, \mathbb{F}_2 \in \Gamma(\mathcal{P})$ *and words* $w_1', u_2'' \in A^*$ *such that:*

(i) $w_1', u_2'' \in \mathcal{L}(A,T)$ *and* $\iota\mathbb{F}_1\iota\mathbb{F}_2 \in \mathcal{L}(A,T)$*;*

(ii) $[\mathbb{E}_1, \mathbb{E}_2]\phi = (w_1'\phi) \cdot [\mathbb{F}_1, \mathbb{F}_2]\phi \cdot (u_2''\phi)$*;*

(iii) the lengths of the words $(\iota\mathbb{F}_1\iota\mathbb{F}_2)\phi, (\tau\mathbb{F}_1\iota\mathbb{F}_2)\phi, (\tau\mathbb{F}_1\tau\mathbb{F}_2)\phi, (\iota\mathbb{F}_1\tau\mathbb{F}_2)\phi$ *are each bounded above by* $2 \times max\{|\alpha| : (\alpha = \beta) \in R \text{ or } (\beta = \alpha) \in R\}$.

Proof. We prove the result in the case that $\mathbb{E}_1$ and $\mathbb{E}_2$ are both positive edges.

The other cases may be dealt with using the same argument. The proof is given for each item as follows:

(i) Let $\mathbb{E}_1 = (w_1, r, +1, w_2)$ and $\mathbb{E}_2 = (u_1, s, +1, u_2)$ where $w_1, w_2, u_1, u_2 \in A^*$ and $r, s \in R$. Let w_1' be the longest prefix of w_1 which is in $\mathcal{L}(A,T)$. Let $w_1 r_{+1} w_2 u_1 s_{+1} u_2'$ be the shortest prefix of $w_1 r_{+1} w_2 u_1 s_{+1} u_2$ which is in $\mathcal{L}(A,T)$ and has $w_1 r_{+1} w_2 u_1 s_{+1}$ as a prefix. Write $w_1 \equiv w_1' w_1''$ and $w_2 \equiv w_2' w_2''$. Since T is left unitary, it follows that $u_2'' \in \mathcal{L}(A,T)$. We then define $\mathbb{F}_1 = (w_1'', r, +1, w_2)$ and $\mathbb{F}_2 = (u_1, s, +1, u_2')$ which are edges in $\Gamma(\mathcal{P})$. Then $\iota\mathbb{F}_1\iota\mathbb{F}_2 \equiv w_1'' r_{+1} w_2 u_1 s_{+1} u_2' \in \mathcal{L}(A,T)$ since T is left unitary, and both w_1' and $w_1' w_1'' r_{+1} w_2 u_1 s_{+1} u_2'$ belong to $\mathcal{L}(A,T)$.

(ii) From part (i) and the definition of paths we have $[\mathbb{E}_1, \mathbb{E}_2] = w_1' \cdot [\mathbb{F}_1, \mathbb{F}_2] \cdot u_2''$. Hence, applying Lemma 4.6.6 to every edge in the path we obtain:

$$[\mathbb{E}_1, \mathbb{E}_2]\phi = (w_1'\phi) \cdot [\mathbb{F}_1, \mathbb{F}_2]\phi \cdot (u_2''\phi).$$

(iii) By the definitions in part (i) and of paths we have the following four vertex words in the path $[\mathbb{F}_1, \mathbb{F}_2]$:

$$\iota\mathbb{F}_1\iota\mathbb{F}_2 \equiv w_1'' r_{+1} w_2 u_1 s_{+1} u_2', \quad \tau\mathbb{F}_1\iota\mathbb{F}_2 \equiv w_1'' r_{-1} w_2 u_1 s_{+1} u_2',$$
$$\tau\mathbb{F}_1\tau\mathbb{F}_2 \equiv w_1'' r_{-1} w_2 u_1 s_{-1} u_2', \quad \iota\mathbb{F}_1\tau\mathbb{F}_2 \equiv w_1'' r_{+1} w_2 u_1 s_{-1} u_2'.$$

As $[\mathbb{E}_1, \mathbb{E}_2]$ does not split, it is immediate from the definitions that $[\mathbb{F}_1, \mathbb{F}_2]$ does not split. Decompose $\iota\mathbb{F}_1\iota\mathbb{F}_2$ into principal factors $\iota\mathbb{F}_1\iota\mathbb{F}_2 \equiv \alpha_1\alpha_2\ldots\alpha_k \equiv w_1'' r_{+1} w_2 u_1 s_{+1} u_2'$ as in the statement of Lemma 4.5.5. We claim that every word α_i $(1 \leq i \leq k)$ contains at least one of the letters from the subword r_{+1} or one of the letters from the subword s_{+1}. Indeed, α_1 contains at least one letter of r_{+1} since w_1'' has no prefix in $\mathcal{L}(A,T)$ by definition of w_1'. Also, α_k contains at least one letter of s_{+1} by definition of u_2'. Finally, no α_j can be a subword of $w_2 u_1$ in the above decomposition since that would imply a split of $[\mathbb{F}_1, \mathbb{F}_2]$ at the word

$\alpha_1\alpha_2\ldots\alpha_j \in \mathcal{L}(A,T)$, contradicting the fact that $[\mathbb{F}_1, \mathbb{F}_2]$ does not split. This proves the claim.

This is illustrated below in Figure 4.6.4 and hence $|(\iota\mathbb{F}_1\iota\mathbb{F}_2)\phi| = k \leq |r_{+1}| + |s_{+1}| \leq 2 \times max\{|\alpha| : (\alpha = \beta) \in R \text{ or } (\beta = \alpha) \in R\}$. Exactly the same argument applies to the other three words $\tau\mathbb{F}_1\iota\mathbb{F}_2$, $\tau\mathbb{F}_1\tau\mathbb{F}_2$ and $\iota\mathbb{F}_1\tau\mathbb{F}_2$, which completes the proof of (iii).

$\square$

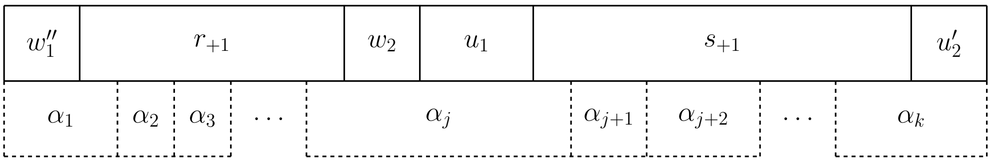


Figure 4.6.4: The decomposition of $\iota\mathbb{F}_1\iota\mathbb{F}_2$ into principal factors in the proof of Lemma 4.6.14 part (iii).

The following lemma summarises the situation for type (5) paths.

Lemma 4.6.15. *Let*

$$N = \{\ [\mathbb{F}_1, \mathbb{F}_2]\phi\ :\ \mathbb{F}_1, \mathbb{F}_2 \in E(\Gamma(\mathcal{P})),\ \iota\mathbb{F}_1\iota\mathbb{F}_2 \in \mathcal{L}(A,T))\ \textit{and}$$

$$max\ (\ |(\iota\mathbb{F}_1\iota\mathbb{F}_2)\phi|, |(\iota\mathbb{F}_1\tau\mathbb{F}_2)\phi|, |(\tau\mathbb{F}_1\iota\mathbb{F}_2)\phi|, |(\tau\mathbb{F}_1\tau\mathbb{F}_2)\phi|\ )$$

$$\leq 2 \times max\ (\ |\alpha| : (\alpha = \beta) \in R\ \textit{or}\ (\beta = \alpha) \in R\ )\ \}.$$

Then N is a finite set of closed paths in $\Gamma(\mathcal{Q})$ such that for any pair of edges $\mathbb{E}_1, \mathbb{E}_2 \in E(\Gamma(\mathcal{P}))$ with $\iota\mathbb{E}_1\iota\mathbb{E}_2 \in \mathcal{L}(A,T)$ we have

$$[\mathbb{E}_1, \mathbb{E}_2]\phi\ \sim_N\ 1_{(\iota\mathbb{E}_1\iota\mathbb{E}_2)\phi}.$$

Proof. If $[\mathbb{E}_1, \mathbb{E}_2]$ splits then $[\mathbb{E}_1, \mathbb{E}_2]\phi\ \sim\ 1_{(\iota\mathbb{E}_1\iota\mathbb{E}_2)\phi}$ by Lemma 4.6.13,

and if $[\mathbb{E}_1, \mathbb{E}_2]$ does not split then $[\mathbb{E}_1, \mathbb{E}_2]\phi\ \sim_N\ 1_{(\iota\mathbb{E}_1\iota\mathbb{E}_2)\phi}$ by Lemma 4.6.14.

So in both cases $[\mathbb{E}_1, \mathbb{E}_2]\phi\ \sim_N\ 1_{(\iota\mathbb{E}_1\iota\mathbb{E}_2)\phi}$ as required. $\square$

Type (6) paths.

We consider the infinite set of closed paths of type (6) where $(w_1 \cdot \mathbb{P} \cdot w_2)\phi$, for $\mathbb{P} \in X$ and $w_1, w_2 \in A^*$ such that $w_1(\iota\mathbb{P})w_2 \in \mathcal{L}(A, T)$. This is illustrated in the following Figure 4.6.5. The notation used in the diagram matches the notation used in the proof of Lemma 4.6.16.

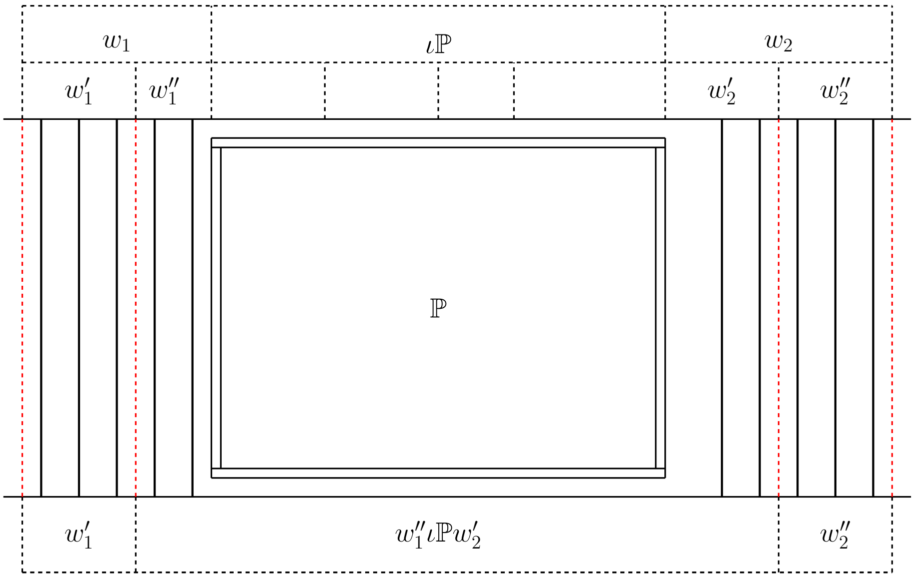


Figure 4.6.5: Type (6) paths

Lemma 4.6.16. *Let $m = max\ \{\ |\iota\mathbb{P}| : \mathbb{P} \in X\ \}$ and then define*

$$K = \{(\alpha \cdot \mathbb{P} \cdot \beta)\phi : \alpha, \beta \in A^*, \mathbb{P} \in X,\ \alpha\iota\mathbb{P}\beta \in \mathcal{L}(A,T)\ \textit{and}\ |\ (\iota(\alpha \cdot \mathbb{P} \cdot \beta))\phi\ | \leq m\}.$$

Then K is a finite set of closed paths in $\Gamma(\mathcal{Q})$ such that for all $w_1, w_2 \in A^*$ and $\mathbb{P} \in X$ such that $w_1\iota\mathbb{P}w_2 \in \mathcal{L}(A,T)$ we have

$$(w_1 \cdot \mathbb{P} \cdot w_2)\phi\ \sim_K\ 1_{v\phi}\ \textit{where}\ v \equiv w_1(\iota\mathbb{P})w_2.$$

Proof. To see that the set K is finite note that if $(\alpha{\cdot}\mathbb{P}{\cdot}\beta)\phi \in K$ then $\iota[(\alpha{\cdot}\mathbb{P}{\cdot}\beta)\phi] \equiv [\iota(\alpha \cdot \mathbb{P} \cdot \beta)]\phi$ so by assumption $|\iota[(\alpha \cdot \mathbb{P} \cdot \beta)\phi]| \leq m$, hence there are finitely many possibilities for the initial vertex of $(\alpha \cdot \mathbb{P} \cdot \beta)\phi$. Also, by definition of ϕ on paths, $(\alpha \cdot \mathbb{P} \cdot \beta)\phi$ is a path with the same number of edges as the path $\mathbb{P}$. But $\mathbb{P} \in X$, which is a finite set of closed paths in $\Gamma(\mathcal{P})$. Hence there is a global bound on the number of edges in the path $(\alpha \cdot \mathbb{P} \cdot \beta)\phi$. Specifically, the length of the path $(\alpha \cdot \mathbb{P} \cdot \beta)\phi$ is no greater than the maximum of the lengths of the paths in X. This shows that for every path $(\alpha \cdot \mathbb{P} \cdot \beta)\phi$ in K there are finitely many possibilities for the initial vertex of this path, and the length of this path is bounded above by the maximum length of the paths in X. Since U is finite, every vertex in $\Gamma(\mathcal{Q})$ is only adjacent to finitely many other vertices. Combining these observations we conclude that K is finite.

For the second part let $w_1, w_2 \in A^*$ and $\mathbb{P} \in X$ such that $w_1(\iota\mathbb{P})w_2 \in \mathcal{L}(A,T)$. Write $w_1(\iota\mathbb{P})w_2 \equiv w_1'w_1''(\iota\mathbb{P})w_2'w_2''$ where w_1' is the longest prefix of w_1 which is in $\mathcal{L}(A,T)$, and $w_1(\iota\mathbb{P})w_2'$ is the shortest prefix of $w_1(\iota\mathbb{P})w_2$ which has $w_1(\iota\mathbb{P})$ as a prefix and is in $\mathcal{L}(A,T)$. Since T is left unitary it follows that $w_2'' \in \mathcal{L}(A,T)$ and $w_1''(\iota\mathbb{P})w_2' \in \mathcal{L}(A,T)$. It follows from these definitions that when we decompose $w_1''(\iota\mathbb{P})w_2'$ into principal factors $w_1''(\iota\mathbb{P})w_2' \equiv \alpha_1\alpha_2\ldots\alpha_k$ as in the statement of Lemma 4.5.5, each of these factors α_i must contain at least one letter from the subword $\iota\mathbb{P}$. Hence $|(w_1''(\iota\mathbb{P})w_2')\phi| \leq |\iota\mathbb{P}| \leq m$ since $\mathbb{P} \in X$. It follows that $(w_1'' \cdot \mathbb{P} \cdot w_2')\phi \in K$. Finally, applying Lemma 4.6.6 we have

$$(w_1 \cdot \mathbb{P} \cdot w_2)\phi = (w_1'\phi) \cdot [(w_1'' \cdot \mathbb{P} \cdot w_2')\phi] \cdot (w_2''\phi) \ \sim_K \ 1_{v\phi}$$

where $v \equiv w_1(\iota\mathbb{P})w_2$, as required. □

Summary for paths of types (4), (5) and (6):

Lemma 4.6.17. *Let W be the finite set of closed paths as defined in Lemma 4.6.11. Let N be the finite set of closed paths as defined in Lemma 4.6.15. Let K be the finite set of closed paths as defined in Lemma 4.6.16.*

Then the homotopy relation generated by $W \cup N \cup K$ is a finite homotopy base for $\Gamma(\mathcal{Q})$.

Proof. By Lemma 4.6.10 Z is an infinite homotopy base for $\Gamma(\mathcal{Q})$. By Lemmas

4.6.11, 4.6.15 and 4.6.16 it follows that for every path $\mathbb{P} \in Z$ we have

$$\mathbb{P} \sim_{W \cup N \cup K} 1_{\iota\mathbb{P}} \text{ in } \Gamma(\mathcal{Q}).$$

It follows that $W \cup N \cup K$ is a finite homotopy base for $\Gamma(\mathcal{Q})$. □

4.6.6 Proof of Theorem 4.1.1

Proof of Theorem 4.1.1. By Lemma 4.6.17 $W \cup N \cup K$ is a finite homotopy base for $\Gamma(\mathcal{Q})$ and the presentation $\langle B \mid Q \rangle$ which defines the subsemigroup T. Thus if the semigroup S has finite derivation type, then the subsemigroup T also has finite derivation type.

This concludes the proof for a left unitary subsemigroup with finite strict right boundary. The proof for a right unitary subsemigroup T with finite strict left boundary is dual. □

4.7 Applications of new FDT theorem

4.7.1 Completely simple semigroups

One possible application is to the Rees matrix semigroups which are isomorphic to completely simple semigroups, particularly those with the conditions we have looked at in the previous chapter. That is, where $S = \mathcal{M}[G; I, \Lambda; P]$, G is a group, I and Λ are finite, P contains entries equal to 1_G and we have a sparse generating set. See Section 3.3 in Chapter 3 for further information on the notation and definitions.

Corollary 4.7.1. *Let $S = \mathcal{M}[G; I, \Lambda; P]$ be a Rees matrix semigroup where G is a group, S is defined by a finite presentation with respect to a sparse generating set and one of two conditions is true:*
(i) $I = \{1\}$ and $\Lambda = \{1, 2, \ldots, n\}$;

(ii) $I = \{1, \ldots, m\}$ *and* $\Lambda = \{1\}$.

If S *has finite derivation type then* G *also has finite derivation type.*

Proof. We give an outline proof.

Part (i): The subgroup H_{11} is isomorphic to the group G. If we consider the elements of S as expressed in the triples (i, g, λ) where $i \in I$, $\lambda \in \Lambda$ and $g \in G$, then it can easily be proved that the subgroup H_{11} is a left unitary subsemigroup of S with finite strict right boundary. The result then follows from Theorem 4.1.1.

Part (ii): In this case it can be proved that the subgroup H_{11} is a right unitary subsemigroup of S with finite strict left boundary. The result then follows. □

Note that the above application already follows from an existing more general result which was proved in [38, Theorem 2].

4.7.2 Completely 0-simple semigroups

Another possible application is to the Rees matrix semigroups which are isomorphic to completely 0-simple semigroups. First we define a Rees matrix semigroup which is isomorphic to a completely 0-simple semigroup.

Definition 4.7.2. [25, Lemma 3.2.2] Let G be a group with identity element e, and let I, Λ be non-empty sets. Let $P = (p_{\lambda j})$ be a $\Lambda \times I$ matrix with entries in the 0-group G^0 $(= G \cup \{0\})$, and suppose that P is *regular*, in the sense that no row or column of P consists entirely of zeros. Formally,

$$(\forall j \in I)(\exists \lambda \in \Lambda)\ p_{\lambda j} \neq 0 \text{ and } (\forall \lambda \in \Lambda)(\exists j \in I)\ p_{\lambda j} \neq 0.$$

Let $S = (I \times G \times \Lambda) \cup \{0\}$, and define a composition on S by

$$(i, a, \lambda)(j, b, \mu) = \begin{cases} (i, ap_{\lambda j}b, \mu) & \text{if } p_{\lambda j} \neq 0, \\ 0 & \text{if } p_{\lambda j} = 0, \end{cases}$$

$$(i, a, \lambda)0 = 0(i, a, \lambda) = 00 = 0.$$

Then S is a completely 0-simple semigroup.

The result that can be deduced using our new Theorem 4.1.1 is as follows:

Theorem 4.7.3. *Let S be a completely 0-simple semigroup with finitely many $\mathcal{R}$- and $\mathcal{L}$-classes. Let G be the unique non-zero maximal subgroup of S. If S has finite derivation type then G has finite derivation type.*

Proof. This theorem can be proved in exactly the same way as [16, Theorem 8.30], combined with applying Theorem 4.1.1.

In summary, let G be a maximal subsemigroup of S. Let T be the union of the group $\mathcal{H}$-classes that intersect the $\mathcal{R}$-class of G. Then it can be proved that the subsemigroup T is right unitary in S and has finite strict left boundary in S. It can also be proved that the subsemigroup G is left unitary in T and has finite strict right boundary in T. □

Note that the above application already follows from an existing more general result which was proved in [38, Theorem 2].

4.8 Potential Future work

4.8.1 Special presentations

As mentioned earlier, Y. Kobayashi shows in his paper [29] that every one-relator monoid has FDT. However, there is an interesting problem relating to special monoids which provides a potential application of Theorem 4.1.1. This is an application that I would like to prove if I had more time. It concerns monoids with what is known as a special presentation, namely $M = \langle A \mid w_1 = 1, \dots, w_k = 1 \rangle$, which prove to have an interesting structure.

<u>Fact:</u> Let M be the monoid defined by the presentation $\langle A \mid w_1 = 1, \dots, w_k = 1 \rangle$ and let $U(M)$ be the group of units of M. Then $U(M)$ is a left unitary subsemigroup of M with a finite strict right boundary.

This fact is not obvious but can be proved using results from the paper [55] by L. Zhang. Once this fact is proved, then by applying Theorem 4.1.1 it should be possible to obtain the following new result:

Conjecture 4.8.1. *Let M be the monoid defined by the presentation $\langle A \mid w_1 = 1, w_2 = 1, \dots, w_k = 1 \rangle$ for some fixed $k \in \mathbb{N}$. Let $U(M)$ be the group of units of M. Then M has FDT if and only if $U(M)$ has FDT.*

Plan of the proof:

($\Leftarrow$) Adapt the argument from the above paper of Y. Kobayashi. (Note that in this paper it is observed that for a one-relator monoid, the group of units $U(M)$ of M has FDT since it is a one-relator group.)

($\Rightarrow$) As described above, prove that $U(M)$ is a left unitary submonoid with strict right boundary in M. The result follows by application of Theorem 4.1.1.

4.8.2 Open questions

(i) The paper [54], discussed in the Introduction, considered the property of being a finite complete rewriting system and proved that this was inherited by large subsemigroups. It is natural to look at the question of whether this property can be inherited by subsemigroups with weaker finiteness conditions. In this chapter we have seen that FDT can be inherited by left (resp. right) unitary subsemigroups with finite strict right (resp. left) boundary. It may be possible to extend this result to subsemigroups with finite left, right or two-sided boundaries. Ultimately, it may be possible to prove that having a FCRS is also inherited by one or both of these types of subsemigroup. At the moment these questions remain open.

(ii) Let S be a regular semigroup defined by a finite complete rewriting system. Let H be a subgroup with finitely many $\mathcal{H}$-classes in its $\mathcal{R}$-class. Here we are looking at a particular $\mathcal{R}$-class and then a specific one of its $\mathcal{H}$-classes. We would like to prove that H is defined by a FCRS but an alternative would be to prove that it has FDT.

4.8.3 Properties beyond FDT

It might be interesting to investigate whether it might be possible to prove analogues of results proved for FDT in this thesis, for other important finiteness properties related to string rewriting systems, see [43].

5

Bicyclic monoid and finitely presented subsemigroups

5.1 Introduction to the bicyclic monoid

In this chapter we will investigate properties of finitely generated subsemigroups of the bicyclic monoid. The bicyclic monoid **B** is defined by the presentation $\langle b, c \mid bc = 1 \rangle$. The first published description of the bicyclic monoid was given by Evgeny Lyapin in 1953. A.H. Clifford and G. Preston claim that whilst working with D. Rees it was independently discovered at some point before 1943. In [10] it is referred to as the simplest member of an extensive class of semigroups known as the bisimple inverse semigroups with identity element. As such it is a very useful semigroup in the theory of simple semigroups.

Two papers [13] and [14] by L. Descalço and N. Ruškuc proved some interesting properties of subsemigroups of the bicyclic monoid. In [14] it is proved that every finitely generated subsemigroup of **B** is finitely presented. This leads us naturally to ask two further questions that are the theme of this research:

- Does every finitely generated subsemigroup of the bicyclic monoid have finite derivation type?
- Does every finitely generated subsemigroup of the bicyclic monoid have a presentation which admits a finite complete rewriting system?

We show that the answer to both of these is positive, that is, every finitely generated subsemigroup of **B** admits a presentation by a finite complete semigroup rewriting system (see Theorem 5.2.1) and consequently also has FDT (see Corollary 5.3.1).

Before proving this theorem we first need to recall some background results on the bicyclic monoid and from the two papers [13] and [14].

5.1.1 Background and definitions

The bicyclic monoid **B** is defined by the presentation $\langle b, c \mid bc = 1\rangle$.

Definition 5.1.1. The natural set of normal forms for **B** is $\{c^i b^j : i, j \geq 0\}$. In fact it is easy to see that $\langle b, c \mid bc = 1\rangle$ is a FCRS and these are the irreducible words with respect to this system. Throughout we use these normal forms to make reference to elements of the bicyclic monoid. The elements of **B** multiply in the following way:

$$c^i b^j c^k b^l = \begin{cases} c^{i-j+k} b^l \text{ if } j \leq k \\ c^i b^{j-k+l} \text{ if } j > k. \end{cases}$$

The bicyclic monoid can be expressed diagrammatically as an infinite square grid, see Figure 5.1.1 below and [25, Equation 1.6.3]. Each square is an element in the monoid and is represented by a word in normal form. This is in fact the egg box picture of the $\mathcal{D}$-class of 1 in **B** [25, Exercise 2]. For more on the bicyclic monoid see [25].

	0	1	2	3
0	1	b	b^2	b^3
1	c	cb	cb^2	cb^3
2	c^2	c^2b	c^2b^2	c^2b^3
3	c^3	c^3b	c^3b^2	c^3b^3

Figure 5.1.1: Bicyclic monoid

5.1.2 Results from previous papers

Next we reproduce those definitions and results from the two papers [13] and [14] which are necessary in order to understand the new results obtained later in this chapter. Of particular note are [13, Theorem 3.1], [14, Theorem 2.1] and [14, Theorem 4.3]. All three theorems are reproduced later for reference. The authors of [13, 14] prove that any subsemigroup of the bicyclic monoid falls into one of five different forms. They go on to prove that all finitely generated subsemigroups are finitely presented. Given the reliance on results from [13] and [14], this chapter is best read with both of these papers to hand.

Define the function $\widehat{\ }: \mathbf{B} \to \mathbf{B}$ by $c^i b^j \mapsto \widehat{c^i b^j} = c^j b^i$. Geometrically $\widehat{\ }$ is a reflection with respect to the diagonal where $D = \{c^i b^i : i \geq 0\}$ is the diagonal. Algebraically this function is an anti-isomorphism of $\mathbf{B}$ i.e. for $x, y \in \mathbf{B}$ then $\widehat{xy} = \widehat{y}\widehat{x}$. Recall that an anti-isomorphism between structured sets A and B is an isomorphism from A to the opposite of B, see Definitions 2.2.2 and 2.2.1.

Define the following subsets of elements of the bicyclic monoid. Let $q, p, i, j \in \mathbb{N}_0$, then

$$D = \{c^i b^i : i \geq 0\} \text{ is the } \textit{diagonal,}$$

$$L_p = \{c^i b^j : 0 \leq j < p\,; i \geq 0\} \text{ is the } \textit{left strip} \text{ (determined by } p\text{)},$$

$$T_{q,p} = \{c^i b^j : q \leq i \leq j < p\} \text{ is a } \textit{triangle.}$$

Note that if $q = p$ then the triangle is an empty set. For $i, m \geq 0$ and $d > 0$ with $I \subseteq \{0, \ldots, m-1\}$ we define the *lines*:

$$\begin{aligned}
\Lambda_i &= \{c^i b^j : j \geq 0\},\\
\Lambda_{i,m,d} &= \{c^i b^j : d \,|\, (j-i), j \geq m\},\\
\Lambda_{I,m,d} &= \bigcup_{i \in I} \Lambda_{i,m,d} = \{c^i b^j : i \in I, d \,|\, (j-i), j \geq m\}.
\end{aligned}$$

For $p \geq 0$, $d > 0$, $r \in [d] = \{0, \ldots, d-1\}$ and $P \subseteq [d]$ we define the *squares*:

$$\begin{aligned}
\Sigma_p &= \{c^i b^j : i, j \geq p\},\\
\Sigma_{p,d,r} &= \{c^{p+r+ud}\, b^{p+r+vd} : u, v \geq 0\},\\
\Sigma_{p,d,P} &= \bigcup_{r \in P} \Sigma_{p,d,r} = \{c^{p+r+ud}\, b^{p+r+vd} : r \in P, u, v \geq 0\}.
\end{aligned}$$

Theorem 5.1.2. *[13, Theorem 3.1] Let S be a subsemigroup of the bicyclic monoid. Then one of the following conditions holds:*

1. *The subsemigroup is a subset of the diagonal; $S \subseteq D$.*

2. *The subsemigroup is a union of a subset of a triangle, a subset of the diagonal above the triangle, a square below the triangle and some lines belonging to a strip determined by the square and the triangle, or the reflection of this union with respect to the diagonal. Formally there exist $q, p \in \mathbb{N}_0$ with $q \leq p$, $d \in \mathbb{N}$, $I \subseteq \{q, \ldots, p-1\}$ with $q \in I$, $P \subseteq \{0, \ldots, d-1\}$ with $0 \in P$, $F_D \subseteq D \cap L_q$, $F \subseteq T_{q,p}$ such that S is one of the following forms:*

   (i) $S = F_D \cup F \cup \Lambda_{I,p,d} \cup \Sigma_{p,d,P}$;

   (ii) $S = F_D \cup \widehat{F} \cup \widehat{\Lambda_{I,p,d}} \cup \Sigma_{p,d,P}$.

3. *There exist $d \in \mathbb{N}$, $I \subseteq \mathbb{N}_0$, $F_D \subseteq D \cap L_{min(I)}$ and sets $S_i \subseteq \Lambda_{i,i,d}$ $(i \in I)$ such that S is one of the following forms:*

   (i) $S = F_D \cup \bigcup_{i \in I} S_i$;

(ii) $S = F_D \cup \bigcup_{i \in I} \widehat{S_i}$;

where each S_i *has the form*

$$S_i = F_i \cup \Lambda_{i,m_i,d}$$

for some $m_i \in \mathbb{N}_0$ *and some finite set* F_i*, and*

$$I = I_0 \cup \{r + ud : r \in R, u \in \mathbb{N}_0, r + ud \geq N\}$$

for some (possibly empty) $R \subseteq \{0, \ldots, d-1\}$*, some* $N \in \mathbb{N}_0$ *and some finite set* $I_0 \subseteq \{0, \ldots, N-1\}$.

Arising from this theorem there are different kinds of subsemigroups which the authors give names to. We list each of these, and identify them with the corresponding part in Theorem 5.1.2 , as follows:

Part 1: We will continue to call this a subset of the diagonal.

Part 2: This corresponds to subsemigroups having elements both above and below the diagonal; we call them *two-sided subsemigroups.* We observe that a subsemigroup defined by 2.(ii) is symmetric to the corresponding subsemigroup given by 2.(i) with respect to the diagonal, and so we can use the anti-isomorphism $\widehat{\ }$ to obtain one from the other.

Part 3: We call *upper subsemigroups* those having all elements above the diagonal and *lower subsemigroups* those having all elements below the diagonal. Part 3. corresponds to upper and lower subsemigroups. Again, 3.(i) and 3.(ii) give subsemigroups symmetric with respect to the diagonal.

Every subsemigroup has one of these forms, on the other hand it is not true that for every choice of parameters there exists a corresponding subsemigroup. In fact the authors of [13] state that they do not have a nice set of conditions that would tell us which choices of parameters yield subsemigroups, and which do not (although some restrictions are implicitly present).

It may assist the reader to reference pictorial examples of the various subsemigroups of $\mathbf{B}$. For a two-sided subsemigroup see [13, Figure 5] and [14, Figure 2]. For an upper subsemigroup see [13, Figure 6] and also see [14, Figure 3] for both an upper and lower subsemigroup. We replicate one such example below.

Figure 5.1.2 is taken from [14, Figure 2] and is an example of a two-sided subsemigroup of $\mathbf{B}$, defined by $d = 3$, $F_D = \{cb\}$, $F = \{c^4b^7\}$, $I = \{4, 5, 7, 8\}$, $p = 10$, $P = \{0, 1\}$.

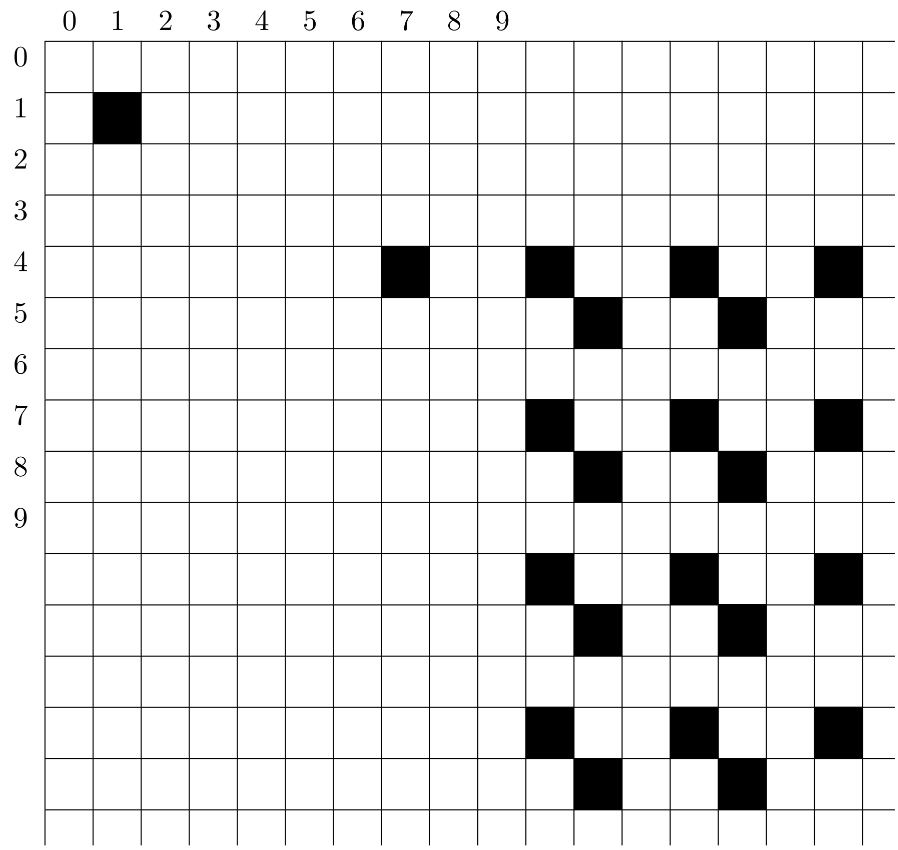


Figure 5.1.2: Bicyclic monoid two-sided semigroup

Lemma 5.1.3. *[13, Lemma 4.7] For any* $p \in \mathbb{N}_0$, $d \in \mathbb{N}$ *and* $I \subseteq \{0, \ldots, p-1\}$, *the set* $\Lambda_{I,p,d}$ *is a subsemigroup.*

Lemma 5.1.4. *[13, Lemma 4.8] Let* $p \in \mathbb{N}_0$, $d \in \mathbb{N}$, $\varnothing \neq I \subseteq \{0, \ldots, p-1\}$,

$\varnothing \neq P \subseteq \{0, \ldots, d-1\}$ and $q = min\{I\}$. The set $H = \Sigma_{p,d,P} \cup \Lambda_{I,p,d}$ is a subsemigroup if and only if

$$I' = \{p + r - ud : r \in P, u \in \mathbb{N}_0, p + r - ud \geq q\} \subseteq I.$$

Theorem 5.1.5. *[14, Theorem 2.1] Let S be a subsemigroup of the bicyclic monoid. Then S is finitely generated if and only if one of the following conditions holds:*

(i) S is a finite diagonal subsemigroup,

(ii) S is a two-sided subsemigroup,

(iii) S is an upper subsemigroup and the set $\{i \in \mathbb{N}_0 : L_i \cap S \neq \varnothing\}$ is finite,

(iv) S is a lower subsemigroup and the set $\{i \in \mathbb{N}_0 : \widehat{L_i} \cap S \neq \varnothing\}$ is finite.

Recall $L_p = \{c^i b^j : 0 \leq j \leq p, i \geq 0\}$ is the left strip determined by $p \geq 0$. This in effect means that there is a finite (non-zero) number of elements within the union of the diagonal and the triangle for cases (iii) and (iv).

Theorem 5.1.6. *[14, Theorem 4.3] All finitely generated subsemigroups of the bicyclic monoid are finitely presented.*

Note that the proof for Theorem 5.1.6 makes use of Lemmas 5.1.3 and 5.1.4 together with Theorem 2.6.18 where the finite small extension is comprised of $F_D \cup F$ for each subsemigroup. A similar approach will be used in the new result in respect of proving we have a finite complete rewriting system.

The next two theorems include useful definitions and facts that are proved within [13] and [14] and are relevant to the new results that we prove later in this chapter. The following results give explicit finite presentations for certain subsemigroups of **B**, and the sets of normal forms for these subsemigroups with respect to these presentations.

Theorem 5.1.7. *[14, Proofs of Theorem 4.3, Lemma 4.4(ii), Lemma 3.3(iii)] Let $p, q \in \mathbb{N}_0$ with $p > q$, $d \in \mathbb{N}$ and let $I \subseteq \{q, \dots, p-1\}$, $P \subseteq \{0, \dots, d-1\}$ such that $0 \in P$. Furthermore, suppose that $T = \Lambda_{I,p,d} \cup \Sigma_{p,d,P}$ is a subsemigroup of the bicyclic monoid.*

Let $Z = \Lambda \cup \{x, y\} \cup \Gamma$ where $\Lambda = \{\lambda_i; i \in I\}$ and $\Gamma = \{\gamma_r; r \in P\}$ be an alphabet.

Let f be a surjective homomorphism from Z^+ to the subsemigroup T such that

$$f : Z^+ \to T,\ \lambda_i \mapsto c^i b^{i+u_i d},\ \gamma_r \mapsto c^{p+r} b^{p+r},\ x \mapsto c^p b^{p+d},\ y \mapsto c^{p+d} b^p$$

where $i + u_i d = min\{i + ud : i + ud \geq p\}$ for $i \in I$.

Let R be the following set of relations in $Z^+ \times Z^+$:

(2) $x = \gamma_0 x$,

(3) $y\gamma_0 = y$,

(4) $\lambda_i \lambda_j = \lambda_i x^{u_j}$ $(i, j \in I)$,

(5) $x\lambda_i = x^{1+u_i}$ $(i \in I)$,

(6) $y\lambda_i = yx^{u_i}$ $(i \in I)$,

(7) $\gamma_r \lambda_i = \gamma_r x^{u_i}$ $(r \in P, i \in I)$,

(8) $xy = \gamma_0$,

(9) $\lambda_i y = \lambda_j$ $(i \in I, u_i > 1, j = p + d - u_i d)$,

(10) $\lambda_i y = \gamma_0$ $(i \in I, u_i = 1)$,

(11) $\gamma_r y = y$ $(r \in P)$,

(12) $x\gamma_r = x$ $(r \in P)$,

(13) $\lambda_i \gamma_r = \lambda_i$ $(i \in I, r \in P, i + u_i d \geq p + r)$,

(14) $\lambda_i \gamma_r = \lambda_j$ $(i \in I, r \in P, i + u_i d < p + r, j = p + r - u_i d)$,

(15) $\gamma_r \gamma_t = \gamma_r$ *(*$r, t \in P$ *and* $r \geq t$*),*

(16) $\gamma_r \gamma_t = \gamma_t$ *(*$r, t \in P$ *and* $r < t$*).*

Then $\langle Z \mid R \rangle$ *is a finite semigroup presentation for* T *and the map* f *induces a well-defined isomorphism from the semigroup defined by* $\langle Z \mid R \rangle$ *to* T *such that*

$$L = \bigcup_{i \in I} \Big(\{ \lambda_i x^u : u \geq 0 \} \Big) \cup \bigcup_{r \in P} \Big(\{ y^v \gamma_r x^u : u, v \geq 0 \} \Big)$$

is a set of unique normal forms for T*.*

Proof. See [14, Proof of Theorem 4.3], [14, Proof of Lemma 4.4(ii)] and also as a consequence of parts of [14, Proof of Lemma 3.3(iii)]. □

Theorem 5.1.8. *[14, Proofs of Theorem 4.3, Lemma 4.4(i), Lemma 3.3(i)] Let* $p, m \in \mathbb{N}_0$ *with* $p \leq m$*,* $d \in \mathbb{N}$ *and the set* $I \subseteq \{0, \ldots, p-1\}$*. Furthermore, suppose that* $T = \Lambda_{I,m,d}$ *is a subsemigroup of the bicyclic monoid.*

Let $i + u_i d = min\{i + ud : i + ud \geq m\}$ *for* $i \in I$*. Fixing* $i_0 \in I$ *and* $u = u_{i_0}$ *we define the alphabet*

$$\Lambda = \bigcup_{i \in I} \{ \lambda(i, 0), \ldots, \lambda(i, u-1) \}$$

and the surjective homomorphism

$$f : \Lambda^+ \to \Lambda_{i,m,d}, \ \lambda(i, j) \mapsto c^i b^{i + (u_i + j)d}.$$

Let R *be the following set of relations in* $\Lambda^+ \times \Lambda^+$

$$\lambda(i, j) \lambda(k, l) = \lambda(i, r) \lambda(i_0, 0)^q$$

where $j + u_k + l = qu + r$*,* $0 \leq r < u$*,* $(i, k \in I;\ j, l \in \{0, \ldots, u-1\})$*.*

Then $\langle \Lambda \mid R \rangle$ *is a finite semigroup presentation for* T *and the map* f *induces a* *well-defined isomorphism from the semigroup defined by* $\langle \Lambda \mid R \rangle$ *to* T *such that*

$$L = \bigcup_{i \in I} \Big(\bigcup_{j=0}^{u-1} \{\lambda(i,j)\lambda(i_0,0)^n : n \geq 0\} \Big)$$

is a set of unique normal forms for T.

Proof. See [14, Proof of Lemma 4.3], [14, Proof of Lemma 4.4(i)] and also as a consequence of parts of [14, Proof of Lemma 3.3(i)]. □

5.2 New research regarding subsemigroups of the bicyclic monoid and the property FCRS

5.2.1 Statement of new theorem

In this section we prove:

Theorem 5.2.1. *Let* $\mathbf{B}$*, defined by the presentation* $\langle b, c \mid bc = 1 \rangle$*, be the bicyclic monoid. Then every finitely generated subsemigroup of* $\mathbf{B}$ *admits a presentation by a finite complete semigroup rewriting system.*

Remark:

Throughout this chapter, for consistency with the two papers [14] and [13], we will be working with subsemigroups of the bicyclic monoid. We will refer to presentations and rewriting systems for semigroups as, for example $\langle Z \mid R \rangle$ and (Z, R) respectively. This means that we will be working with an alphabet Z, words in Z^+ and relations $R \subseteq Z^+ \times Z^+$. In theorems, for extra clarity, we will be specific and call these semigroup presentations and semigroup rewriting systems. We could have chosen to work in terms of submonoids, rather than subsemigroups. Note that there is no major difference between these two classes in the sense that for every finitely generated subsemigroup S of $\mathbf{B}$, if S is not a monoid then $S \cup \{1\}$ is a finitely generated submonoid of $\mathbf{B}$. In fact the result in this section could easily be adapted to prove that every finitely generated submonoid of **B** admits a finite complete monoid rewriting system.

5.2.2 Outline of proof

We prove Theorem 5.2.1 by considering each of the five different forms of finitely generated subsemigroups of the bicyclic monoid **B** as defined in Theorems 5.1.2 and 5.1.5. For our purposes we identify them as follows:

(a) A finite subset of the diagonal, which we identify as $S_1 \subseteq D$.

(b) A two-sided subsemigroup and its reflection in the diagonal, identified as $S_2 = F_D \cup F \cup \Lambda_{I,p,d} \cup \Sigma_{p,d,P}$ and $S_3 = F_D \cup \widehat{F} \cup \widehat{\Lambda_{I,p,d}} \cup \Sigma_{p,d,P}$ respectively. Without loss of generality we need only consider one of these in detail and we choose S_2.

(c) An upper or lower subsemigroup, which we identify as $S_4 = F_D \cup F \cup \Lambda_{I,p,d}$ and $S_5 = F_D \cup \widehat{F} \cup \widehat{\Lambda_{I,p,d}}$ respectively. Without loss of generality we need only consider one of these in detail and we choose S_4.

The proof builds on the presentations defined in Theorems 5.1.7 and 5.1.8. For each form of subsemigroup we define a presentation which is a finite complete rewriting system. The anti-isomorphism property of subsemigroups S_3 and S_5 (with S_2 and S_4 respectively) is used for their proofs. By Theorem 5.1.5 we will have considered all forms of finitely generated subsemigroups and so the proof will be complete.

5.2.3 Finite subset of the diagonal of B

Lemma 5.2.2. *Let S_1 be a subsemigroup of the bicyclic monoid which is a finite subset of the diagonal as defined above in this section. Then S_1 admits a presentation which is a finite complete rewriting system.*

Proof. It is known that a finite semigroup admits a finite presentation which is a finite complete rewriting system, see [24, Section 12.3, Page 423]. As S_1 is finite, the result follows. $\square$

5.2.4 Two-sided subsemigroup of B

Let $S_2 = F_D \cup F \cup \Lambda_{I,p,d} \cup \Sigma_{p,d,P}$ be a two-sided subsemigroup of the bicyclic monoid **B** as defined in Theorem 5.1.2 and Lemma 5.1.4. Note that the sets F_D and F are finite and that $\Lambda_{I,p,d} \cup \Sigma_{p,d,P}$ is a subsemigroup by Lemma 5.1.4. Let

$$U_2 := S_2 \setminus (F_D \cup F) = \Lambda_{I,p,d} \cup \Sigma_{p,d,P}.$$

Let Z be an alphabet and f the mapping $f : Z^+ \to U_2$, both as defined in Theorem 5.1.7. Define the following set of string rewriting rules R' in $Z^+ \times Z^+$:

(2)$'$ $\gamma_0 x \to x$,

(3)$'$ $y\gamma_0 \to y$,

(4)$'$ $\lambda_i\lambda_j \to \lambda_i x^{u_j}$ $(i, j \in I)$,

(5)$'$ $x\lambda_i \to x^{1+u_i}$ $(i \in I)$,

(6)$'$ $y\lambda_i \to yx^{u_i}$ $(i \in I)$,

(7)$'$ $\gamma_r\lambda_i \to \gamma_r x^{u_i}$ $(r \in P, i \in I)$,

(8)$'$ $xy \to \gamma_0$,

(9)$'$ $\lambda_i y \to \lambda_j$ $(i \in I, u_i > 1, j = p + d - u_i d)$,

(10)$'$ $\lambda_i y \to \gamma_0$ $(i \in I, u_i = 1)$,

(11)$'$ $\gamma_r y \to y$ $(r \in P)$,

(12)$'$ $x\gamma_r \to x$ $(r \in P)$,

$(13)'$ $\lambda_i\gamma_r \to \lambda_i$ $(i \in I, r \in P, i + u_i d \geq p + r)$,

$(14)'$ $\lambda_i\gamma_r \to \lambda_j$ $(i \in I, r \in P, i + u_i d < p + r, j = p + r - u_i d)$,

$(15)'$ $\gamma_r\gamma_t \to \gamma_r$ $(r, t \in P$ and $r \geq t)$,

$(16)'$ $\gamma_r\gamma_t \to \gamma_t$ $(r, t \in P$ and $r < t)$.

Note that the rules R' are the relations from R but expressed in terms of a rewrite rule i.e. they have a direction for rewriting. The direction has been chosen to enable the string rewriting system to have the noetherian property, which we will prove later. For our purposes we cannot use the set of normal forms defined as L in Theorem 5.1.7, instead we define the following set $L' \subseteq Z^+$, which we will prove is the set of unique normal forms for the presentation $\langle Z \mid R' \rangle$. Let

$$L' = \bigcup_{i \in I} \big(\{\lambda_i x^u : u \geq 0\}\big) \cup \bigcup_{r \in P \setminus \{0\}} \big(\{y^v \gamma_r x^u : u, v \geq 0\}\big) \cup \{y^v x^u : u, v \geq 0\} \cup \{\gamma_0\}.$$

The definitions and notation above relating to S_2, U_2, f, Z, R', L' will stay in force for the remainder of Subsection 5.2.4. The rest of this subsection will be devoted to proving that (Z, R') is a complete semigroup rewriting system defining the subsemigroup U_2 and L' is the set of irreducible words with respect to this presentation. As a consequence we will conclude with a proof that there exists a finite semigroup presentation for S_2 which is a finite complete semigroup rewriting system.

Lemma 5.2.3. *The subsemigroup U_2 is defined by the finite presentation $\langle Z \mid R' \rangle$.*

Proof. By Theorem 5.1.7, $\langle Z \mid R \rangle$ is a finite presentation for the semigroup U_2. It can be seen that the congruence classes induced on Z^+ by $\overset{*}{\leftrightarrow}_R$ and $\overset{*}{\leftrightarrow}_{R'}$ are the same since R and R' are equal when considered as unordered sets of defining relations. Therefore $\langle Z \mid R' \rangle$ is a finite presentation for U_2. $\square$

Next we aim to prove that the rewriting system (Z, R') is a finite complete rewriting system.

Lemma 5.2.4. *The set $L' \subseteq Z^+$ is the set of normal forms where each word uniquely represents an element in U_2 as defined by the presentation $\langle Z \mid R' \rangle$.*

Proof. We show that each element in U_2 has a unique representative in L' by first establishing a bijection between the sets L' and L. Then we prove that if $l' \in L'$ maps to $l \in L$, then they represent the same element in $\mathbf{B}$. We start by defining subsets of L' and L to facilitate the definition of a map between them.

Let $L' = A' \cup B' \cup C' \cup D'$ where A', B', C' and D' are disjoint subsets such that

$$\begin{aligned}
A' &= \{\lambda_i x^u : i \in I, u \geq 0\} \\
B' &= \{y^v \gamma_r x^u : r \in P \setminus \{0\}; u, v \geq 0\} \\
C' &= \{y^v x^u : u, v \geq 0 \text{ and not both zero}\} \\
D' &= \{\gamma_0\}.
\end{aligned}$$

Let $L = A \cup B \cup C \cup D$ where A, B, C and D are disjoint subsets such that

$$\begin{aligned}
A &= \{\lambda_j x^m : j \in I, m \geq 0\} \\
B &= \{y^n \gamma_q x^m : q \in P \setminus \{0\}; m, n \geq 0\} \\
C &= \{y^n \gamma_0 x^m : m, n \geq 0 \text{ and not both zero}\} \\
D &= \{\gamma_0\}.
\end{aligned}$$

Define the mapping $g : L' \to L$ where $l' \in L'$ and

$$\begin{aligned}
&\text{if } l' \in A' \text{ then } \lambda_i x^u \mapsto \lambda_i x^u \in A; \\
&\text{if } l' \in B' \text{ then } y^v \gamma_r x^u \mapsto y^v \gamma_r x^u \in B; \\
&\text{if } l' \in C' \text{ then } y^v x^u \mapsto y^v \gamma_0 x^u \in C; \\
&\text{if } l' \in D' \text{ then } \gamma_0 \mapsto \gamma_0 \in D.
\end{aligned}$$

It can be seen that for each of the subsets of L' and L, the mapping g is a bijection. As $L' = A' \cup B' \cup C' \cup D'$ and $L = A \cup B \cup C \cup D$, then g is a bijection between the sets L' and L. The second part of the proof will show that for every $l' \in L'$ we have $(l')f = \big((l')g\big)f$ in $\mathbf{B}$. Again we look at each subset of L' in turn and use Definition 5.1.1 for multiplying the elements in $\mathbf{B}$.

(a) Let $l' \equiv \lambda_i x^u \in A'$ with $i \in I$ and $u \geq 0$

then $(l')g \equiv \lambda_i x^u \;\Rightarrow (l')f \equiv \big((l')g\big)f$.

(b) Let $l' \equiv y^v \gamma_r x^u \in B'$ with $r \in P \setminus \{0\}$ and $u, v \geq 0$

then $(l')g \equiv y^v \gamma_r x^u \;\Rightarrow (l')f \equiv \big((l')g\big)f$.

(c) Let $l' \equiv y^v x^u \in C'$ with $u, v \geq 0$ and not both zero

then $\big((l')g\big)f = (y^v \gamma_0 x^u)f = (c^{p+d}b^p)^v\,(c^p b^p)\,(c^p b^{p+d})^u$;

if $v > 0 \Rightarrow \big((l')g\big)f = (c^{p+d}b^p)^{v-1}\,(c^{p+d}b^p)\,(c^p b^p)\,(c^p b^{p+d})^u$

$\Rightarrow \big((l')g\big)f = (c^{p+d}b^p)^{v-1}\,(c^{p+d}b^p)\,(c^p b^{p+d})^u$

$\Rightarrow \big((l')g\big)f = (c^{p+d}b^p)^v(c^p b^{p+d})^u \;=\; (l')f$;

if $u > 0 \Rightarrow \big((l')g\big)f = (c^{p+d}b^p)^v\,(c^p b^p)\,(c^p b^{p+d})\,(c^p b^{p+d})^{u-1}$

$\Rightarrow \big((l')g\big)f = (c^{p+d}b^p)^v\,(c^p b^{p+d})\,(c^p b^{p+d})^{u-1}$

$\Rightarrow \big((l')g\big)f = (c^{p+d}b^p)^v(c^p b^{p+d})^u \;=\; (l')f$.

(d) Let $l' \equiv \gamma_0 \in D'$

then $(l')g \equiv \gamma_0 \;\Rightarrow (l')f \equiv \big((l')g\big)f$.

Thus for every element $l' \in L'$ we have $(l')f = \big((l')g\big)f$ in $\mathbf{B}$ i.e. elements that map one to the other using the mapping g represent the same element in $\mathbf{B}$. We have shown that g is a bijection between the sets L' and L. By Theorem 5.1.7, L is a unique set of normal forms for the subsemigroup U_2 defined by the presentation $\langle Z \mid R \rangle$. Therefore L' is the unique set of normal forms for the subsemigroup U_2 defined by the presentation $\langle Z \mid R' \rangle$. □

Lemma 5.2.5. *The set L' is the set of irreducible words in Z^+ which uniquely represents the elements in U_2 with respect to the string rewriting system (Z, R').*

Proof. By Lemma 5.2.4 we have shown that the set L' uniquely represents every element in U_2 as defined by the presentation $\langle Z \mid R' \rangle$. It remains to show that $L' = \mathrm{IRR}(R')$.

We show that the words in L' cannot be reduced (or rewritten) any further by the string rewriting system R'. We take each subset of L' as follows:

A' Substrings of $\lambda_i x^u$ are $\lambda_i x$ and xx. These do not appear as the left hand side of any of the rewrite rules in R'. Therefore this word cannot be reduced any further.

B' Substrings of $y^v \gamma_r x^u$ with $r \neq 0$, $u, v \geq 0$ are yy, $y\gamma_r$, $\gamma_r x$, xx and $y\gamma_r x$. These do not appear as the left hand side of any of the rewrite rules in R'. Therefore this word cannot be reduced any further.

C' Substrings of $y^v x^u$ with $u, v \geq 0$ but not both zero are yy, yx and xx. These do not appear as the left hand side of any of the rewrite rules in R'. Therefore this word cannot be reduced any further.

D' The letter γ_0 does not appear as the left hand side of any of the rewrite rules in R'. Therefore this word cannot be reduced any further.

Thus all the words in L' are irreducible under R' and so $L' \subseteq \mathrm{IRR}(R')$. It remains to prove that $\mathrm{IRR}(R') \subseteq L'$ i.e. if any word, say $w \in Z^+$ is irreducible then it must be in the set L'. The proof is by induction on the length of the word w.

<u>Induction statement</u>

Let $w \in Z^+$ be irreducible and let $w_1 \equiv wz$ where $z \in Z$ such that $|w_1| = |w| + 1$. Next apply the rewrite rules R' to w_1 such that $w_1 \xrightarrow{*}_{R'} w_1' \in \mathrm{IRR}(R')$, then $w_1' \in L'$.

<u>Base case</u>

If $|w| = 0$ then $w_1 \in Z$ and w_1 is irreducible and we have $w \in L'$.

Let $w \in Z^+$ be irreducible with $|w| \geq 1$ and let $z \in Z$. It follows from the inductive hypothesis that $w_1' \in L'$. There are now several cases to consider:

(i) Let $w \equiv \lambda_i$ then

if $w_1 \equiv \lambda_i\lambda_j \xrightarrow{(4)'} \lambda_i x^{u_j} \in L'$;

if $w_1 \equiv \lambda_i x \in L'$;

if $w_1 \equiv \lambda_i y \xrightarrow{(10)'} \gamma_0 \in L'$;

if $w_1 \equiv \lambda_i\gamma_r \xrightarrow{(13)'} \lambda_i \in L'$, or $w_1 \xrightarrow{(14)'} \lambda_j \in L'$.

(ii) Let $w \equiv x^u$ with $u \geq 1$ then

if $w_1 \equiv x^u\lambda_i \xrightarrow{(5)'} x^{u+u_i} \in L'$;

if $w_1 \equiv x^u x \in L'$;

if $w_1 \equiv x^u y$ and $u = 1$ $w_1 \xrightarrow{(8)'} \gamma_0 \in L'$, but if $u > 1$ $w_1 \xrightarrow{(12)'} x^{u-1} \in L'$;

if $w_1 \equiv x^u\gamma_r \xrightarrow{(12)'} x^u \in L'$.

(iii) Let $w \equiv \lambda_i x^u$ with $u \geq 1$ then

if $w_1 \equiv \lambda_i x^u\lambda_i \xrightarrow{(5)'} \lambda_i x^{u+u_i} \in L'$;

if $w_1 \equiv \lambda_i x^u x \in L'$;

if $w_1 \equiv \lambda_i x^u y$ and $u = 1$ $w_1 \xrightarrow{(8,13/14)'} \lambda_{i/j} \in L'$, but if $u > 1$ $w_1 \xrightarrow{(8,12)'} \lambda_i x^{u-1} \in L'$;

if $w_1 \equiv \lambda_i x^u\gamma_r \xrightarrow{(12)'} \lambda_i x^u \in L'$.

(iv) Let $w \equiv y^v$ with $v \geq 1$ then

if $w_1 \equiv y^v\lambda_i \xrightarrow{(6)'} y^v x^{u_i} \in L'$;

if $w_1 \equiv y^v x \in L'$;

if $w_1 \equiv y^v y \in L'$;

if $w_1 \equiv y^v\gamma_r$ and $r \neq 0$ $w_1 \in L'$ but if $r = 0$ $w_1 \xrightarrow{(3)'} y^v \in L'$.

(v) Let $w \equiv \gamma_r$ with $r \neq 0$ then

if $w_1 \equiv \gamma_r\lambda_i \xrightarrow{(7)'} \gamma_r x^{u_i} \in L'$;

if $w_1 \equiv \gamma_r x \in L'$;

if $w_1 \equiv \gamma_r y \xrightarrow{(11)'} y \in L'$;

if $w_1 \equiv \gamma_r\gamma_t$ and $r \geq t$ $w_1 \xrightarrow{(15)'} \gamma_r \in L'$ but if $r < t$ $w_1 \xrightarrow{(16)'} \gamma_t \in L'$.

(vi) Let $w \equiv y^v\gamma_r$ with $r \neq 0$ and $v \geq 1$ then if $w_1 \equiv y^v\gamma_r\lambda_i \xrightarrow{(7)'} y^v\gamma_r x^{u_i} \in L'$;

if $w_1 \equiv y^v\gamma_r x \in L'$;

if $w_1 \equiv y^v\gamma_r y \xrightarrow{(11)'} y^{v+1} \in L'$;

if $w_1 \equiv y^v\gamma_r\gamma_t$ and $r \geq t$ $w_1 \xrightarrow{(15)'} y^v\gamma_r \in L'$, but if $r < t$ $w_1 \xrightarrow{(16)'} y^v\gamma_t \in L'$.

(vii) Let $w \equiv \gamma_r x^u$ with $r \neq 0$ and $u \geq 1$ then

if $w_1 \equiv \gamma_r x^u\lambda_i \xrightarrow{(5)'} \gamma_r x^{u+u_i} \in L'$;

if $w_1 \equiv \gamma_r x^u x \in L'$;

if $w_1 \equiv \gamma_r x^u y$ and $u = 1$ $w_1 \xrightarrow{(8,15)'} \gamma_r \in L'$, but if $u > 1$ $w_1 \xrightarrow{(8,12)'} \gamma_r x^{u-1} \in L'$;

if $w_1 \equiv \gamma_r x^u\gamma_t \xrightarrow{(12)'} \gamma_r x^u \in L'$.

(viii) Let $w \equiv y^v\gamma_r x^u$ with $r \neq 0$ and $u, v \geq 1$ then

if $w_1 \equiv y^v\gamma_r x^u\lambda_i \xrightarrow{(5)'} y^v\gamma_r x^{u+u_i} \in L'$;

if $w_1 \equiv y^v\gamma_r x^u x \in L'$;

if $w_1 \equiv y^v\gamma_r x^u y$ and $u = 1$ $w_1 \xrightarrow{(8,15)'} y^v\gamma_r \in L'$, but if $u > 1$ $w_1 \xrightarrow{(8,12)'} y^v\gamma_r x^{u-1} \in L'$;

if $w_1 \equiv y^v\gamma_r x^u\gamma_t \xrightarrow{(12)'} y^v\gamma_r x^u \in L'$.

(ix) Let $w \equiv y^v x^u$ with $u, v \geq 1$ then

if $w_1 \equiv y^v x^u\lambda_i \xrightarrow{(5)'} y^v x^{u+u_i} \in L'$;

if $w_1 \equiv y^v x^u x \in L'$;

if $w_1 \equiv y^v x^u y$ and $u = 1$ $w_1 \xrightarrow{(8,3)'} y^v \in L'$, but if $u > 1$ $w_1 \xrightarrow{(3,12)'} y^v x^{u-1} \in L'$;

if $w_1 \equiv y^v x^u\gamma_r \xrightarrow{(12)'} y^v x^u \in L'$.

(x) Let $w \equiv \gamma_0$ then

if $w_1 \equiv \gamma_0\lambda_i \xrightarrow{(7,2)'} x^{u_i} \in L'$;

if $w_1 \equiv \gamma_0 x \xrightarrow{(2)'} x \in L'$;

if $w_1 \equiv \gamma_0 y \xrightarrow{(11)'} y \in L'$;

if $w_1 \equiv \gamma_0\gamma_r \xrightarrow{(16)'} \gamma_r \in L'$.

We have shown that our induction statement is true for $|w| = 0$ and also for $|w_1| = |w| + 1$ where $|w| \geq 1$, therefore $\mathrm{IRR}(R') \subseteq L'$. As we have already shown that $L' \subseteq \mathrm{IRR}(R')$, then we must have $L' = \mathrm{IRR}(R')$. □

Next we aim to prove that the rewriting system (Z, R') is noetherian. If we look at the rewrite rules R' we can gain some understanding of how they act when they are applied to words in Z^+. A single application of a rewrite rule to a word in Z^+ can have the following affect:

- It can reduce the number of letters in a word which are equal to one from the set Λ.
- It can reduce the length of the word.
- It can remove a letter which is equal to one from the set Λ and at the same time add letters to the word. The increase in the length of the word is determined by the value of the letter from the set Λ which is removed, say $\lambda_i \in \Lambda$, and is at most the value of $u_i - 1$ which is fixed for each $i \in I$ and bounded as $0 \leq u_i - 1 \leq p - 1$.
- It cannot introduce a new letter which is equal to one from the set Λ.

Now we define an ordering on words in Z^+ before we later go on to prove that it is a reduction ordering on Z^+, induced by R'.

Definition 5.2.6. Let $w \in Z^+$ be an arbitrary word. Define w_Λ to be the number of letters in the word w which are equal to one of the letters from the set Λ. Recall that $|w|$ is the number of letters in the word w. Let σ be the mapping defined by

$$\sigma : Z^+ \to \mathbb{N}_0 \times \mathbb{N}_0 \text{ where } w \mapsto (w_\Lambda, |w|).$$

Let $a, b \in Z^+$ be arbitrary words with $a_1, a_2, b_1, b_2 \in \mathbb{N}_0$, such that $a\sigma = (a_1, a_2)$ and $b\sigma = (b_1, b_2)$. Then we say that $b <_\sigma a$ if one of the following is true:
(i) $b_1 < a_1$,
(ii) $b_1 = a_1$ and $b_2 < a_2$.

Lemma 5.2.7. *The rewriting system (Z, R') is noetherian.*

Proof. It is useful at this point to recall the requirements for a set of rewrite rules to be noetherian. It suffices to find a reduction ordering $>$ on Z^+ which is an *admissible*, *well-founded partial ordering* such that $u > v$ holds for each rule $(u, v) \in R'$. See Theorem 2.6.7, Definition 2.6.1 and Definition 2.6.3. We aim to show that this is true for the ordering $<_\sigma$, see Definition 5.2.6.

First we prove that $<_\sigma$ is *admissible*, see Definition 2.6.1. Let $\alpha, \beta, w_2 \in Z^*$, $w_1 \in Z^+$ such that $w_2 <_\sigma w_1$, $s_1 = \alpha w_1 \beta$ and $s_2 = \alpha w_2 \beta$. Then
$s_1\sigma = (\alpha w_1 \beta)\sigma = (\alpha_\Lambda + (w_1)_\Lambda + \beta_\Lambda\ ,\ |\alpha| + |w_1| + |\beta|)$ and
$s_2\sigma = (\alpha w_2 \beta)\sigma = (\alpha_\Lambda + (w_2)_\Lambda + \beta_\Lambda\ ,\ |\alpha| + |w_2| + |\beta|)$.

If $w_2 <_\sigma w_1$ with

(i) $(w_2)_\Lambda < (w_1)_\Lambda$ then
$\alpha_\Lambda + (w_2)_\Lambda + \beta_\Lambda \ < \ \alpha_\Lambda + (w_1)_\Lambda + \beta_\Lambda$ and so $s_2 <_\sigma s_1$;

(ii) $(w_2)_\Lambda \ = \ (w_1)_\Lambda$ and $|w_2| \ < \ |w_1|$ then
$\alpha_\Lambda + (w_2)_\Lambda + \beta_\Lambda \ = \ \alpha_\Lambda + (w_1)_\Lambda + \beta_\Lambda$ and
$|\alpha| + |w_2| + |\beta| \ < \ |\alpha| + |w_1| + |\beta|$ and so $s_2 <_\sigma s_1$.

Thus in each case $s_2 <_\sigma s_1$ and therefore $<_\sigma$ is an admissible ordering.

To prove that $<_\sigma$ is a well-ordering (see Definition 2.6.3) we need to prove two properties. Firstly that $<_\sigma$ is a strict partial ordering. Secondly that it is well-founded i.e. there is no infinite descending chain with
$\ldots w_{h+1} <_\sigma w_h <_\sigma \ldots <_\sigma w_2 <_\sigma w_1$ where $w_i \in Z^+$.

Let $x, y, z \in Z^+$. The ordering $<_\sigma$ is irreflexive as $x \not<_\sigma x$ by definition of $<_\sigma$. If $y <_\sigma x$ then by definition of $<_\sigma$ it is not true that $x <_\sigma y$, therefore $<_\sigma$ is anti-symmetric. If $x <_\sigma y$ and $y <_\sigma z$, then by definition of $<_\sigma$, we have $x <_\sigma z$ and therefore $<_\sigma$ is transitive. So by Definition (a) 2.6.1 $<_\sigma$ is a strict partial ordering on Z^+.

Let $w \in Z^+$ and $x_1, x_2 \in \mathbb{N}_0$ such that $w\sigma = (x_1, x_2)$ then $<_\sigma$ is an ordering on the Cartesian product (x_1, x_2) where each of x_1, x_2 is ordered by $<_{sh}$. By Lemma 2.6.2 the ordering $<_{sh}$ is well-founded, therefore by Lemma 2.6.6, $<_\sigma$ is well-founded.

It remains to prove that for every $(u, v) \in R'$ we have $v <_\sigma u$. We examine each rule in turn:

$(2)'$ $(\gamma_0 x,\ x)$ where $u\sigma = (0, 2)$ and $v\sigma = (0, 1) \Rightarrow v <_\sigma u$.
$(3)'$ $(y\gamma_0,\ y)$ where $u\sigma = (0, 2)$ and $v\sigma = (0, 1) \Rightarrow v <_\sigma u$.
$(4)'$ $(\lambda_i\lambda_j,\ \lambda_i x^{u_j})$ where $u\sigma = (2, 2)$ and $v\sigma = (1, 1 + u_j) \Rightarrow v <_\sigma u$.
$(5)'$ $(x\lambda_i,\ x^{1+u_i})$ where $u\sigma = (1, 2)$ and $v\sigma = (0, 1 + u_i) \Rightarrow v <_\sigma u$.
$(6)'$ $(y\lambda_i,\ yx^{u_i})$ where $u\sigma = (1, 2)$ and $v\sigma = (0, 1 + u_i) \Rightarrow v <_\sigma u$.
$(7)'$ $(\gamma_r\lambda_i,\ \gamma_r x^{u_i})$ where $u\sigma = (1, 2)$ and $v\sigma = (0, 1 + u_i) \Rightarrow v <_\sigma u$.
$(8)'$ $(xy,\ \gamma_0)$ where $u\sigma = (0, 2)$ and $v\sigma = (0, 1) \Rightarrow v <_\sigma u$.
$(9)'$ $(\lambda_i y,\ \lambda_j)$ where $u\sigma = (1, 2)$ and $v\sigma = (1, 1) \Rightarrow v <_\sigma u$.
$(10)'$ $(\lambda_i y,\ \gamma_0)$ where $u\sigma = (1, 2)$ and $v\sigma = (0, 1) \Rightarrow v <_\sigma u$.
$(11)'$ $(\gamma_r y,\ y)$ where $u\sigma = (0, 2)$ and $v\sigma = (0, 1) \Rightarrow v <_\sigma u$.
$(12)'$ $(x\gamma_r,\ x)$ where $u\sigma = (0, 2)$ and $v\sigma = (0, 1) \Rightarrow v <_\sigma u$.
$(13)'$ $(\lambda_i\gamma_r,\ \lambda_i)$ where $u\sigma = (1, 2)$ and $v\sigma = (1, 1) \Rightarrow v <_\sigma u$.
$(14)'$ $(\lambda_i\gamma_r,\ \lambda_j)$ where $u\sigma = (1, 2)$ and $v\sigma = (1, 1) \Rightarrow v <_\sigma u$.
$(15)'$ $(\gamma_r\gamma_t,\ \gamma_r)$ where $u\sigma = (0, 2)$ and $v\sigma = (0, 1) \Rightarrow v <_\sigma u$.
$(16)'$ $(\gamma_r\gamma_t,\ \gamma_t)$ where $u\sigma = (0, 2)$ and $v\sigma = (0, 1) \Rightarrow v <_\sigma u$.

Hence for all $(u, v) \in R'$ we have $v <_\sigma u$. Therefore we have defined a reduction ordering $<_\sigma$ on Z^+ which is compatible with R' and so the rewriting system (Z, R') is noetherian. □

Theorem 5.2.8. *Let $U_2 = \Lambda_{I,p,d} \cup \Sigma_{p,d,P}$ be a subsemigroup of the bicyclic monoid, Z an alphabet, L' the set of words in Z^+ and R' the set of rewrite rules in $Z^+ \times Z^+$, all as defined above.*

Then (Z, R') is a finite complete semigroup rewriting system for U_2.

Proof. By Lemma 5.2.3, $\langle Z \mid R' \rangle$ is a finite semigroup presentation for U_2. By Lemma 5.2.7, the rewriting system (Z, R') is noetherian. By Lemma 5.2.5, L' is the unique set of normal forms with $L' = \mathrm{IRR}(R')$. Then by Theorem 2.6.14 and Lemma 2.6.10 we can say that (Z, R') is a finite complete semigroup rewriting system for U_2. $\square$

Lemma 5.2.9. *Let $S_2 = F_D \cup F \cup \Lambda_{I,p,d} \cup \Sigma_{p,d,P}$ be the two sided subsemigroup of the bicyclic monoid as defined above.*

Then there exists a finite presentation for S_2 which is a finite complete rewriting system.

Proof. Let $U_2 = \Lambda_{I,p,d} \cup \Sigma_{p,d,P}$. By Theorem 5.2.8, $\langle Z \mid R' \rangle$ is a finite presentation for U_2 which is a finite complete rewriting system. By definition, $S_2 \setminus U_2 = F_D \cup F$ is a finite set. The result follows from this and Corollary 2.6.21. $\square$

5.2.5 Upper subsemigroup of B

Let $S_4 = F_D \cup F \cup \Lambda_{I,p,d}$ be an upper subsemigroup of the bicyclic monoid **B** as defined in Theorem 5.1.2 and Lemma 5.1.3. Note that the sets F_D and F are finite and that $\Lambda_{I,p,d}$ is a subsemigroup by Lemma 5.1.3. Let

$$U_4 := S_4 \setminus F_D \cup F = \Lambda_{I,p,d}.$$

Let Λ be an alphabet, f the homomorphism $f : \Lambda^+ \to U_4$ and $L \subseteq \Lambda^+$ the unique set of normal forms, all as defined in Theorem 5.1.8.

Define the following set of string rewriting rules R' in $\Lambda^+ \times \Lambda^+$:

$$\lambda(i,j)\lambda(k,l) \to \lambda(i,r)\lambda(i_0,0)^q \quad \text{if } \lambda(k,l) \not\equiv \lambda(i_0,0);$$

$$\text{where } j + u_k + l = qu + r, 0 \le r < u, (i,k \in I; j,l \in \{0,\dots,u-1\}).$$

Note that the rules R' are the relations from R but expressed in terms of a rewrite rule. However, we have the extra condition i.e. $\lambda(k,l) \not\equiv \lambda(i_0,0)$. This change is in order to achieve the noetherian property for the rewrite system (Λ, R') which we will prove later.

The definitions and notation above relating to $S_4, U_4, f, \Lambda, L, R'$ will stay in force for the remainder of Subsection 5.2.5. The rest of this subsection will be devoted to proving that (Λ, R') is a complete semigroup rewriting system defining the subsemigroup U_4 and L is the set of irreducible words with respect to this presentation. As a consequence we will conclude with a proof that there exists a finite semigroup presentation for S_4 which is a finite complete semigroup rewriting system.

Lemma 5.2.10. *The subsemigroup U_4 is defined by the finite presentation $\langle \Lambda \mid R' \rangle$.*

Proof. By Theorem 5.1.8, $\langle \Lambda \mid R \rangle$ is a finite presentation for $\Lambda_{I,p,d} = U_4$. We aim to prove that $\langle \Lambda \mid R' \rangle$ is equivalent to $\langle \Lambda \mid R \rangle$.

We are considering the semigroup presentation $\langle \Lambda \mid R' \rangle$ and therefore we can compare $\Lambda^+/ \overset{*}{\leftrightarrow}_{R'}$ with $\Lambda^+/ \overset{*}{\leftrightarrow}_{R}$ in terms of their congruence classes in that these represent the elements in the subsemigroup U_4. Most relations R and R' are equal when considered as unordered sets of defining relations with one exception. As such, we only have one potential difference, which is the exclusion of $\lambda(i,j)\lambda(i_0,0) = \lambda(i,r)\lambda(i_0,0)^q$ from the set R'. So, if we have $\lambda(i,j)\lambda(i_0,0)$ on the left hand side under the relations R, then we can calculate what the right hand side would be. To find r and q we use the following:

$$j + u_k + l = qu + r,\ 0 \leq r < u \text{ and } i,k \in I; j,l \in \{0,\ldots,u-1\}.$$

In this case as $k = i_0$ we have $u_k = u$ and also $l = 0$ which gives us
$j + u + 0 = qu + r$. Also, we know that $j, r \in \{0,\ldots,u-1\}$ and so we can deduce

that as j and r are both strictly less than u then $q = 1$ and it follows that $j = r$.

So, in applying the relations in R we find that our calculations would result in the relation $\lambda(i,j)\lambda(i_0,0) \to \lambda(i,j)\lambda(i_0,0)$, which is clearly redundant and is the relation which we have excluded from the set R'. This means that we have simply removed a redundant relation and therefore by Theorem 2.8.2, the congruence classes for $\Lambda^+/\overset{*}{\leftrightarrow}_{R'}$ and $\Lambda^+/\overset{*}{\leftrightarrow}_{R}$ are the same. Hence $\langle \Lambda \mid R' \rangle$ is a finite presentation for U_4. $\square$

Next we aim to prove that the rewriting system (Λ, R') is a finite complete rewriting system.

Lemma 5.2.11. *The set $L \subseteq \Lambda^+$ is a set where each word uniquely represents an element in U_4 as defined by the presentation $\langle \Lambda \mid R' \rangle$.*

Proof. By Theorem 5.1.8 $L \subseteq \Lambda^+$ is the unique set of normal forms for the presentation $\langle \Lambda \mid R \rangle$ and therefore they represent every element in the subsemigroup U_4. In Lemma 5.2.10 it was proved that the congruence classes $\Lambda^+/\overset{*}{\leftrightarrow}_{R'}$ and $\Lambda^+/\overset{*}{\leftrightarrow}_{R}$ are the same. Therefore $L \subseteq \Lambda^+$ is a set where each word uniquely represents an element in U_4 as defined by the presentation $\langle \Lambda \mid R' \rangle$. $\square$

Lemma 5.2.12. *The set L is the set of irreducible words in Λ^+ which uniquely represents the elements in U_4 with respect to the string rewriting system (Λ, R').*

Proof. By Lemma 5.2.11 we have shown that L uniquely represents every element in U_4 as defined by the presentation $\langle \Lambda \mid R' \rangle$. It remains to prove that $L = \mathrm{IRR}(R')$.

Let l be an arbitrary word in L and now we look for any substring of letters in l that appear within the left hand side of any rewrite rule in R'. If such a substring does not exist then the word is irreducible by the rewrite rules in R'. Possible substrings are as follows:

(a) $\lambda(i,j)\lambda(i_0,0)$ which is specifically excluded from the rewrite rules as the second letter is $\lambda(i_0,0)$ and so this substring will not be rewritten.

(b) $\lambda(i_0,0)\lambda(i_0,0)$ as above for (a).

(c) $\lambda(i,j)$ there is no rewrite rule for a single letter.

(d) $\lambda(i_0,0)$ there is no rewrite rule for a single letter.

Thus all the words in L are irreducible under R' and therefore $L \subseteq \mathrm{IRR}(R')$. It remains to prove that $\mathrm{IRR}(R') \subseteq L$, that is, if any word, say $w \in \Lambda^+$ is irreducible then it must be in the set L. The proof is by induction on the length of the word w.

Induction statement

Let $w \in \Lambda^+$ be irreducible and let $w_1 \equiv wa$ where $a \in \Lambda$ such that $|w_1| = |w|+1$. Next apply the rewrite rules R' to w_1 such that $w_1 \xrightarrow{*}_{R'} w_1' \in \mathrm{IRR}(R')$, then $w_1' \in L$.

Base case

If $|w| = 0$ then $w_1 \in \Lambda$ and w_1 is irreducible, and we have $w \in L$.

Let $w \in \Lambda^+$ be irreducible with $|w| \geq 1$ and let $a \in \Lambda$. It follows from the inductive hypothesis that $w_1' \in L'$. There are now several cases to consider, note that $i_0, i, k \in I$ and $j, l, r, s, t \in \{0, \ldots, u-1\}$:

(i) Let $w_1 \equiv \lambda(i_0,0)\lambda(k,l)$ and
if $\lambda(k,l) \equiv \lambda(i_0,0)$ then $w_1 \in L$;
if $\lambda(k,l) \not\equiv \lambda(i_0,0)$ then $w_1 \to \lambda(i_0,r)\lambda(i_0,0)^q \in L$.

(ii) Let $w_1 \equiv \lambda(i,j)\lambda(k,l)$ with $\lambda(i,j) \not\equiv \lambda(i_0,0)$ and
if $\lambda(k,l) \equiv \lambda(i_0,0)$ then $w_1 \in L$;
if $\lambda(k,l) \not\equiv \lambda(i_0,0)$ then $w_1 \to \lambda(i,r)\lambda(i_0,0)^q \in L$.

(iii) Let $w_1 \equiv \lambda(i,j)\lambda(i_0,0)\lambda(k,l)$ with $\lambda(i,j) \not\equiv \lambda(i_0,0)$ and
if $\lambda(k,l) \equiv \lambda(i_0,0)$ then $w_1 \in L$;

if $\lambda(k,l) \not\equiv \lambda(i_0,0)$ then $w_1 \to \lambda(i,j)\lambda(i_0,s)\lambda(i_0,0)^q \to \lambda(i,r)\lambda(i_0,0)^{q_1} \in L$.

(iv) Let $w_1 \equiv \lambda(i,j)\lambda(i_0,0)^q\lambda(k,l)$ with $\lambda(i,j) \not\equiv \lambda(i_0,0)$, $q > 1$ and
if $\lambda(k,l) \equiv \lambda(i_0,0)$ then $w_1 \in L$;
if $\lambda(k,l) \not\equiv \lambda(i_0,0)$ then
$w_1 \to \lambda(i,j)\lambda(i_0,0)^{q-1}\lambda(i_0,s)\lambda(i_0,0)^{q_1} \xrightarrow{*} \lambda(i,r)\lambda(i_0,0)^{q'} \in L$.

We have shown that our induction statement is true for $|w| = 0$ and also for $|w_1| = |w| + 1$ where $|w| \geq 1$, therefore $\mathrm{IRR}(R') \subseteq L$. As we have already shown that $L \subseteq \mathrm{IRR}(R')$, then we must have $L = \mathrm{IRR}(R')$. $\square$

Next we aim to prove that the rewriting system (Λ, R') is noetherian. If we look at the rewrite rules R' we can gain some understanding of how they act when they are applied to words in Λ^+. A single application of a rewrite rule to a word can have the following affect

- It can move a letter which is not equal to $\lambda(i_0,0)$ left one place and at the same time the number of letters in the word may increase, but only as a result of the introduction of new words that are equal to $\lambda(i_0,0)$.
- It can remove a letter which is not equal to $\lambda(i_0,0)$ and replace it with zero, one or many letters that are equal to $\lambda(i_0,0)$.
- It cannot introduce a new letter which is not equal to $\lambda(i_0,0)$.

Now we define an ordering on words in Λ^+ before we later go on to prove that it is a reduction ordering on Λ^+, induced by R'.

Definition 5.2.13. Let $w \in \Lambda^+$ be an arbitrary word. Let ϕ be the mapping of words in Λ^+ to a finite sequence of integers defined by

$$\phi : \Lambda^+ \to \bigcup_{m\in\mathbb{N}} (\mathbb{N})^m \text{ with } w\phi \mapsto (n_1, n_2, \ldots, n_m)$$

where each $n_h \in \{n_1, \ldots, n_m\}$ represents the numbered position of a letter $\lambda_j \in \Lambda$ within the word w, where $\lambda_j \not\equiv \lambda(i_0,0)$ and the numbering of the position for n_h

is from left to right in the word.

For example, let $w_1 \equiv \lambda(i_0,0)\lambda(i,j)\lambda(k,l)\lambda(i_0,0)\lambda(i_0,0)\lambda(g,h)$ where we have $\lambda(i,j),\lambda(k,l),\lambda(g,h) \not\equiv \lambda(i_0,0)$ then $w_1\phi = (2,3,6)$.

Let $a,b \in \Lambda^+$ such that $b\phi = (n_1,\ldots,n_k)$ and $a\phi = (m_1,\ldots,m_l)$, then we say $b <_\phi a$ if one of the following is true:

(i) $k < l$

(ii) $k = l$ and the smallest value for $p \in \{1,2,\ldots,k\}$ where $n_p \neq m_p$ is such that $n_p < m_p$.

Here are some examples in order to clarify the above Definition 5.2.13:

(1) Let $w_1\phi = (1,3,7)$ and $w_2 = (2,10)$ then $w_2 <_\phi w_1$.

(2) Let $w_1\phi = (2,4,5,9,12)$ and $w_2 = (2,3,5,7,20)$ then $w_2 <_\phi w_1$.

(3) Let $w_1\phi = (1,4)$ and $w_2 = (1,4)$ then $w_2 \not<_\phi w_1$ and $w_1 \not<_\phi w_2$.

Lemma 5.2.14. *The rewriting system (Λ, R') is noetherian.*

Proof. We will show that $<_\phi$, as per Definition 5.2.13, is an admissible, well-founded partial ordering on Λ^+ and then go on to use this to prove that the rewriting system (Λ, R') is noetherian.

We prove that $<_\phi$ is an admissible ordering. Let $s \in \Lambda^+$, $\alpha,\beta,t \in \Lambda^*$, $a \equiv \alpha s\beta$ and $b \equiv \alpha t\beta$ such that $t <_\phi s$. Let $\alpha\phi = (c_1,\ldots,c_d)$, $\beta\phi = (e_1,\ldots,e_f)$, $s\phi = (g_1,\ldots,g_h)$ and $t\phi = (m_1,\ldots,m_n)$. By definition of $<_\phi$ we have either

(i) $n < h$ or

(ii) $n = h$ and the smallest value for $p \in \{1,\ldots,n\}$ where $m_p \neq g_p$ is such that $m_p < g_p$.

Now we consider $a\phi$ and $b\phi$:

$a\phi = (c_1,\ldots,c_d,g_1+|\alpha|,\ldots,g_h+|\alpha|,e_1+|\alpha|+|s|,\ldots,e_f+|\alpha|+|s|)$ and

$b\phi = (c_1,\ldots,c_d,m_1+|\alpha|,\ldots,m_n+|\alpha|,e_1+|\alpha|+|t|,\ldots,e_f+|\alpha|+|t|)$.

Note that the number of terms in $a\phi$ is $d+h+f$ and in $b\phi$ is $d+n+f$.

If $t <_\phi s$ and (i) $n < h$ then when we consider the number of terms in $b\phi$ and $a\phi$ we have $d + n + f < d + h + f$ and so $b <_\phi a$.
If $t <_\phi s$ and (ii) then when we consider the number of terms in $b\phi$ and $a\phi$, we have $d + n + f = d + h + f$ and so we look at the value of the terms. By (ii) the smallest value for $p \in \{1, \dots, n\}$ where $m_p + |\alpha| \neq g_p + |\alpha|$ is such that $m_p + |\alpha| < g_p + |\alpha|$ and so $b <_\phi a$.

Thus in each case $b <_\phi a$ and therefore $<_\phi$ is an admissible ordering on Λ^+.

To prove that $<_\phi$ is a well-ordering we need to prove two properties. Firstly that $<_\phi$ is a strict partial ordering on Λ^+. Secondly that it is well-founded i.e. that there is no infinite chain of rewriting under R'.

Let $x, y, z \in \Lambda^+$. The ordering $<_\phi$ is irreflexive as $x \not<_\phi x$ by definition of $<_\phi$. If $y <_\phi x$ then by definition of $<_\phi$ it is not true that $x <_\phi y$ and therefore $<_\phi$ is anti-symmetric. If $x <_\phi y$ and $y <_\phi z$, then by definition of $<_\phi$, we have $x <_\phi z$ and therefore $<_\phi$ is transitive. So by Definition (a) 2.6.1 $<_\phi$ is a strict partial ordering on Λ^+.

Let $w_1, w_2 \in \Lambda^+$, $w_1\phi = (a_1, \dots, a_n)$, $w_2\phi = (b_1, \dots, b_m)$ such that $w_2 <_\phi w_1$. When considering if $<_\phi$ is true, then we have two values that we use. The first one is the number of entries within the brackets for $w_1\phi$ and $w_2\phi$. This is an ordering of the form $(\mathbb{N}_0, <)$ which is well-founded. The second value is only considered in the case where $w_1\phi$ and $w_2\phi$ have the same number of terms i.e. $n = m$. Then we compare the values within the brackets i.e. each a_i and b_i pair, looking for the first difference. Without loss of generality we can assume that $a_i = b_i$ for all $i \in \{1, \dots, x-1\}$ and that $b_x < a_x$. This is equivalent to an ordering of the form $(\mathbb{N}_0, <)$ which is well-founded. So we have an ordering $<_\phi$ which is essentially a $(\mathbb{N}_0 \times \mathbb{N}_0, <_k)$ where $<_k = (<, <)_{lex}$ and $(<, <)_{lex}$ is defined as in Example 2.6.5 (ii). Therefore by Lemma 2.6.6 $<_\phi$ is well-founded.

It remains to prove that for every $(s, t) \in R'$ we have $t <_\phi s$. We consider the possible values for $s\phi$ and $t\phi$ for each of the potential rewriting outcomes. There are two main cases to consider:

Case 1:

Let $s \equiv \lambda(i_0,0)\lambda(k,l)$ with $\lambda(k,l) \not\equiv \lambda(i_0,0)$ and so $t \equiv \lambda(i_0,r)\lambda(i_0,0)^q$. We have two sub cases, one where $r = 0$ and one where $r \neq 0$:

(i) If $t \equiv \lambda(i_0,0)\lambda(i_0,0)^q$ then $t\phi = ()$ and $s\phi = (2)$, therefore $t <_\phi s$.

(ii) If $t \equiv \lambda(i_0,r)\lambda(i_0,0)^q$ then $t\phi = (1)$ and $s\phi = (2)$, therefore $t <_\phi s$.

Case 2:

Let $s \equiv \lambda(i,j)\lambda(k,l)$ where $\lambda(i,j) \not\equiv \lambda(i_0,0) \not\equiv \lambda(k,l)$ and so $t \equiv \lambda(i,r)\lambda(i_0,0)^q$. We have two sub cases, one where $\lambda(i,r) \equiv \lambda(i_0,0)$ and one where $\lambda(i,r) \not\equiv \lambda(i_0,0)$:

(i) If $t \equiv \lambda(i_0,0)\lambda(i_0,0)^q$, then $t\phi = ()$ and $s\phi = (1,2)$, therefore $t <_\phi s$.

(ii) If $t \equiv \lambda(i,r)\lambda(i_0,0)^q$, then $t\phi = (1)$ and $s\phi = (1,2)$, therefore $t <_\phi s$.

To summarise, for every $(s,t) \in R'$ we have $t <_\phi s$. Now we can say that we have defined a reduction ordering $<_\phi$ on Λ^+ which is compatible with R' and therefore the rewriting system (Λ, R') is noetherian. □

Theorem 5.2.15. *Let $U_4 = \Lambda_{I,p,d}$ be a subsemigroup of the bicyclic monoid, Λ an alphabet, L the set of words in Λ^+ and R' the set of rewrite rules in $\Lambda^+ \times \Lambda^+$, all as defined above.*

Then (Λ, R') is a finite complete semigroup rewriting system for U_4.

Proof. By Lemma 5.2.10, $\langle \Lambda \mid R' \rangle$ is a finite presentation for U_4. By Lemma 5.2.14, the rewriting system (Λ, R') is noetherian. By Lemma 5.2.12, L is the unique set of normal forms with $L = \mathrm{IRR}(R')$. Then by Theorem 2.6.14 and Lemma 2.6.10 we can say that (Λ, R') is a finite complete semigroup rewriting system for U_4. □

Lemma 5.2.16. *Let $S_4 = F_D \cup F \cup \Lambda_{I,p,d}$ be an upper subsemigroup of the bicyclic monoid as defined above.*

Then there exists a finite presentation for S_4 which is a finite complete rewriting *system.*

Proof. Let $U_4 = \Lambda_{I,p,d}$. By Theorem 5.2.15, $\langle Z \mid R' \rangle$ is a finite presentation for U_4 which is a finite complete rewriting system. By definition, $S_4 \setminus U_4 = F_D \cup F$ is a finite set. The result follows from this and Corollary 2.6.21. □

5.2.6 Proof of new Theorem 5.2.1

Proof. **Returning to the proof of Theorem 5.2.1.**

By Theorem 5.1.2, if S is a subsemigroup of the bicyclic monoid $\mathbf{B}$, then it takes one of five different forms. By Theorem 5.1.5, there are five forms of subsemigroup which are finitely generated. We refer the reader to previous sections and to [13] and [14] for full definitions. We consider each of the five finitely generated subsemigroups in turn:

(i) A finite subset of the diagonal, which we define as $S_1 \subseteq D$. By Lemma 5.2.2, there exists a presentation for S_1 which is a finite complete semigroup rewriting system.

(ii) A two-sided subsemigroup, which is defined as $S_2 = F_D \cup F \cup \Lambda_{I,p,d} \cup \Sigma_{p,d,P}$. By Lemma 5.2.9, there exists a presentation for S_2 which is a finite complete semigroup rewriting system.

(iii) A two-sided subsemigroup which is defined as $S_3 = F_D \cup \widehat{F} \cup \widehat{\Lambda_{I,p,d}} \cup \Sigma_{p,d,P}$. By Lemma 5.2.9 and the anti-isomorphism of S_3 with S_2, then there exists a presentation for S_3 which is a finite complete semigroup rewriting system.

(iv) An upper subsemigroup, which is defined as $S_4 = F_D \cup F \cup \Lambda_{I,p,d}$. By Lemma 5.2.16, there exists a presentation for S_4 which is a finite complete semigroup rewriting system.

(v) A lower subsemigroup, which is defined as $S_5 = F_D \cup \widehat{F} \cup \widehat{\Lambda_{I,p,d}}$. By Lemma 5.2.16 and the anti-isomorphism of S_5 with S_4 then there exists a presentation for S_5 which is a finite complete semigroup rewriting system.

This completes the proof of Theorem 5.2.1. □

5.3 Further applications

Corollary 5.3.1. *Every finitely generated subsemigroup of the bicyclic monoid has finite derivation type.*

Proof. By Theorem 5.2.1 every finitely generated subsemigroup of the bicyclic monoid admits a presentation which is a finite complete rewriting system. Then by Theorem 4.4.2 and Definition 4.4.4, every finitely generated subsemigroup of the bicyclic monoid has FDT. □

5.4 Potential future work

It would be interesting to identify other monoids, or classes of monoids, with the property that all their finitely generated submonoids admit a finite complete rewriting system (like the bicyclic monoid does). In particular, we do not yet know whether the plactic monoid has this property.

6

Plactic monoid submonoids and homogeneous presentations

6.1 Introduction to the plactic monoid

6.1.1 Plactic monoid definitions and notation

This section includes the relevant definitions and properties of the plactic monoid, some of which have been reproduced from [34, Chapter 5]. Further information and more detail can be found in this work. Also included are definitions for Young diagrams and tableau taken from [15] and [4], where further information can be found. Note that there can be two different conventions for a Young diagram and here we will describe one.

The following definition of the plactic monoid uses the classic Knuth relations.

Definition 6.1.1. [34, Def 5.2.2] Let $n \in \mathbb{N}$. Let A be the finite ordered alphabet $\{1 < 2 < \ldots < n\}$. Let R be the set of Knuth relations:
$\{(xzy, zxy) : x \leq y < z\} \cup \{(yxz, yzx) : x < y \leq z\}$ where $x, y, z \in A$.
Then the *plactic monoid* P_n of rank n is the monoid defined by the presentation $\langle A \mid R \rangle$.

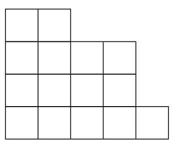
In studying the plactic monoid we will use a combinatorial structure known as a Young tableau. In order to describe this structure we first make some definitions.

A *row* is a non-decreasing word in A^* i.e. a word $\alpha \equiv \alpha_1 \ldots \alpha_k$, where $\alpha_i \in A$, in which $\alpha_i \leq \alpha_{i+1}$ for all $i = 1, \ldots, k-1$. Let $\alpha \equiv \alpha_1 \ldots \alpha_k$ and $\beta \equiv \beta_1 \ldots \beta_l$ (where $\alpha_i, \beta_i \in A$) be rows. The row α *dominates* the row β, denoted $\alpha \triangleright \beta$ if $k \leq l$ and $\alpha_i > \beta_i$ for all $i = 1, \ldots, k$.

Any word $w \in A^*$ has a decomposition as a product of rows of maximal length $w \equiv \alpha^{(1)} \ldots \alpha^{(k)}$, where $\alpha^{(1)}$ is the longest prefix of w that is a row, then $\alpha^{(2)}$ is the longest prefix of the remaining suffix which is a row, and so on. Such a word w is a *tableau* if $\alpha^{(i)} \triangleright \alpha^{(i+1)}$ for all $i = 1, \ldots, k-1$. It is usual to write tableaux in planar form with the rows placed in order of domination and left-justified i.e. as a Young tableau, which we will see shortly.

A *column* is a strictly decreasing word in A^* i.e. a word $\alpha \equiv \alpha_1 \ldots \alpha_l$, where $\alpha_i \in A$, in which $\alpha_{i+1} < \alpha_i$ for all $i = 1, \ldots, l-1$.

A *Young diagram* (also called a Ferrers diagram, particularly when represented using dots) is a finite collection of boxes, or cells, arranged in left-justified rows, with longer rows appearing below shorter rows such that there is a weakly increasing number of boxes in each row. Listing the number of boxes in each row gives a partition of the integer n that is the total number of boxes. Conversely, every partition of n corresponds to a Young diagram. For example, the partition of 15 into $5 + 4 + 4 + 2$ corresponds to the Young diagram:



We write partitions in the form $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)$ where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$. So the partition above would be written as $(5, 4, 4, 2)$. A *Young tableau*, or simply a *tableau*, of shape λ is a *filling* of a Young diagram of the partition λ with positive integers such that the entries are:

(i) weakly increasing across each row;

(ii) strictly decreasing down each column.

The above filling of a Young diagram will result in a *semistandard* tableau. A *standard tableau* is a tableau in which the entries are numbered from 1 to n, each occurring once. We will be working with semistandard tableau when considering the plactic monoid. Entries of tableaux can also be taken from any alphabet (totally ordered set), but positive integers are more usual. The following example is a tableau of shape $(5, 4, 4, 2)$:

$$\begin{array}{|c|c|c|c|c|}
\cline{1-2}
7 & 8 \\
\cline{1-4}
4 & 6 & 7 & 7 \\
\cline{1-4}
2 & 2 & 3 & 4 \\
\hline
1 & 1 & 1 & 3 & 5 \\
\hline
\end{array}$$

The *column-reading* of a tableau is the word obtained by reading the tableau down each column in turn, with the columns ordered left-to-right. Dually, the *row-reading* of a tableau is the word obtained by reading the tableau along each row in turn from left-to-right, with the rows ordered from top-to-bottom. Thus, the row-reading of the above example tableau would result in the word 78 4677 2234 11135. Note that this is a *tableau word* (in the sense defined above). Indeed, it is immediate from the definitions that every tableau word gives rise to a semistandard tableau, and conversely the row reading of any semistandard tableau is a tableau word.

The importance of tableau words is that every word in A^* is equal in the plactic monoid P_n to a unique tableau word (see Theorem 6.1.5 below). We use $P(w)$ to denote the unique tableau word such that $w = P(w)$ in the plactic monoid. The word $P(w)$ can be computed from w using Schensted's algorithm. Next we introduce Schensted's algorithm which is a process to determine the tableau for any given word. Here we use it to obtain a tableau word.

Definition 6.1.2. Let $P : A^* \to A^*$ be the map defined by induction on word length in the following way. We define $P(\epsilon) \equiv \epsilon$ (where ϵ is the empty word) and $P(a) \equiv a$ for every letter $a \in A$. Then for any word $w \in A^*$ with $|w| > 1$ we define $P(w) \equiv P(P(w')\gamma)$ where $w \equiv w'\gamma$, $\gamma \in A$ and $P(P(w')\gamma)$ is computed by applying Schensted's algorithm (see Definition 6.1.3 below) to the tableau word

$P(w')$ with the letter γ.

The following definition and additional paragraph of further explanation have been taken from [4]. Let $w \in A^*$ be a tableau word and let $\gamma \in A$. The unique tableau word $P(w\gamma)$ can be computed via Schensted's algorithm [34] which we recall here:

Definition 6.1.3. [4, Algorithm 2.2] [Shensted's algorithm]

Input: A tableau word w with rows $\alpha^{(1)}, \ldots, \alpha^{(k)}$ and a letter $\gamma \in A$.

Output: The unique tableau word $P(w\gamma)$ equal to $w\gamma$ in P_n.

Method:

Case 1 If $\alpha^{(k)}\gamma$ is a row, the result is $P(w\gamma) \equiv \alpha^{(1)} \ldots \alpha^{(k)}\gamma$.

Case 2 If $\alpha^{(k)}\gamma$ is not a row, the result is determined as follows: Suppose $\alpha^{(k)} \equiv \alpha_1 \ldots \alpha_l$ where $\alpha_i \in A$ for all $i \in \{1, \ldots, l\}$ and let j be minimal such that $\alpha_j > \gamma$. Then the result is $P(w\gamma) = P(\alpha^{(1)} \ldots \alpha^{(k-1)}\alpha_j)\alpha'^{(k)}$ where $\alpha'^{(k)} \equiv \alpha_1 \ldots \alpha_{j-1}\gamma\alpha_{j+1} \ldots \alpha_l$.

Notice that in Case 2, the algorithm replaces α_j by γ in the lowest row and recursively right-multiplies by α_j the tableau formed by all rows except the lowest. This is referred to as *bumping* α_j to a higher row. When α_j is bumped, it will be inserted into the row above either in the same column or in some column further to the left. This happens because columns are strictly decreasing from top to bottom, so either the cell above α_j contains some symbol δ greater than α_j, or α_j is the topmost element of its column. In the former case, α_j will be inserted so as to replace the leftmost symbol greater than α_j, which must either be to the left of δ or δ itself, since rows are non-decreasing from left to right. In the latter case, α_j will be appended to the end of the row above and so will be placed either in the same column or further left.

The following example illustrates the stages of using Schensted's algorithm and shows how it can be applied to compute a tableau word from a given word, starting from scratch. Note that the tableau word can be obtained by row-reading of the Young tableau.

Example 6.1.4. Let P_5 be the plactic monoid on the ordered alphabet $A = \{1 < 2 < \ldots < 5\}$ and $w \equiv 213543 \in A^+$. Then the tableau word $P(w)$ is computed using Schensted's algorithm as follows:

$$\begin{array}{l} 2 \end{array} \quad \begin{array}{l} 2 \\ 1 \end{array} \quad \begin{array}{ll} 2 & \\ 1 & 3 \end{array} \quad \begin{array}{lll} 2 & & \\ 1 & 3 & 5 \end{array} \quad \begin{array}{lll} 2 & 5 & \\ 1 & 3 & 4 \end{array} \quad \begin{array}{lll} 5 & & \\ 2 & 4 & \\ 1 & 3 & 3 \end{array}$$

so that

$$P(213543) \equiv 524133.$$

Note that the rows of this tableau word are 5, 24 and 133.

This brings us to the important fact which relates tableaux to the plactic monoid. Let P_n be the plactic monoid on the alphabet $A = \{1, 2, \ldots, n\}$ and let $u, v \in A^*$. We introduce an equivalence relation $\sim$ on A^* defined by $u \sim v$ if and only if $P(u) \equiv P(v)$.

Theorem 6.1.5. *[34, Theorem 5.2.5] The equivalence $\sim$ coincides with the plactic congruence. In particular, each plactic class contains exactly one tableau word.*

Consider the plactic monoid P_n, let $w_1, w_2 \in A^*$ represent the same element in P_n i.e. $w_1 = w_2$. Then $P(w_1) \equiv P(w_2)$. Note this also means that $|w_1| = |w_2|$ and both words contain exactly the same letters from A but in a different order.

As a further example, we can use Schensted's algorithm to see the equality in the Knuth relations used to define the plactic monoid see Definition 6.1.1.

Let $x < y < z$ then

$$P(xzy) \equiv P(zxy) \; \begin{array}{ll} z & \\ x & y \end{array} \qquad \text{and} \qquad P(yxz) \equiv P(yzx) \; \begin{array}{ll} y & \\ x & z \end{array}$$

Let $x < y$ then

$$P(xyx) \equiv P(yxx) \quad \begin{array}{|c|c}\hline y & \\ \hline x & \multicolumn{1}{c|}{x} \\ \hline \end{array} \qquad \text{and} \qquad P(yxy) \equiv P(yyx) \quad \begin{array}{|c|c}\hline y & \\ \hline x & \multicolumn{1}{c|}{y} \\ \hline \end{array}$$

To illustrate how useful it is to find the tableau form of a word we take an arbitrary word $w = 132541$ in the plactic monoid P_5. The diagram below (Figure 6.1.1) shows relations between certain words, all of which represent the same element in P_5. In fact the diagram is a graph showing part of a connected component of the Squier graph. Recall we encountered these in a previous chapter.

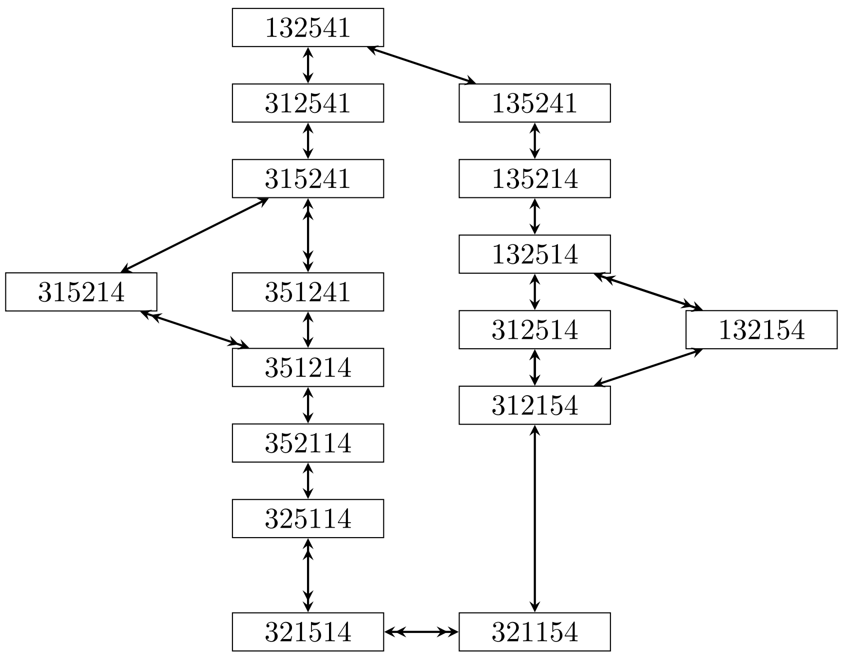


Figure 6.1.1: Notes for the diagram. The boxes represent words in A^*, all of which represent the same element in P_5. A double headed arrow depicts an application of a Knuth relation which can be applied in both directions between two words. In a few cases it is possible to apply a relation to separate sub strings within the words and this case is represented by two double headed arrows. The diagram is not exhaustive and shows only a selection of words which could represent this particular element.

Alternatively, by applying Schensted's algorithm to say $w_1 \equiv 132541$ and $w_2 \equiv 312154$, we can see that $w_1 =_{P_5} w_2$.

$$P(w_1) \equiv P(132541) \quad \begin{array}{|c|}\hline 1 \\ \hline \end{array} \quad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline \end{array} \quad \begin{array}{|c|c}\hline 3 & \\ \hline 1 & \multicolumn{1}{c|}{2} \\ \hline \end{array} \quad \begin{array}{|c|cc}\hline 3 & & \\ \hline 1 & \multicolumn{1}{c|}{2} & \multicolumn{1}{c|}{5} \\ \hline \end{array} \quad \begin{array}{|c|c|c}\hline 3 & 5 & \\ \hline 1 & 2 & \multicolumn{1}{c|}{4} \\ \hline \end{array} \quad \begin{array}{|c|cc}\hline 3 & & \\ \hline 2 & \multicolumn{1}{c|}{5} & \\ \hline 1 & \multicolumn{1}{c|}{1} & \multicolumn{1}{c|}{4} \\ \hline \end{array}$$

and

$$P(w_2) \equiv P(312154) \quad \begin{array}{|c|}\hline 3 \\ \hline\end{array} \quad \begin{array}{|c|}\hline 3 \\ \hline 1 \\ \hline\end{array} \quad \begin{array}{|c|c|}\cline{1-1} 3 & \multicolumn{1}{c}{} \\ \hline 1 & 2 \\ \hline\end{array} \quad \begin{array}{|c|c|}\cline{1-1} 3 & \multicolumn{1}{c}{} \\ \cline{1-1} 2 & \multicolumn{1}{c}{} \\ \hline 1 & 1 \\ \hline\end{array} \quad \begin{array}{|c|c|c|}\cline{1-1} 3 & \multicolumn{2}{c}{} \\ \cline{1-1} 2 & \multicolumn{2}{c}{} \\ \hline 1 & 1 & 5 \\ \hline\end{array} \quad \begin{array}{|c|c|c|}\cline{1-1} 3 & \multicolumn{2}{c}{} \\ \cline{1-2} 2 & 5 & \multicolumn{1}{c}{} \\ \hline 1 & 1 & 4 \\ \hline\end{array}$$

Similarly, if we have $w_3 \in A^*$ where $w_3 \equiv 113254$ and applying Schensted's algorithm we get:

$$P(w_3) \equiv P(113254) \quad \begin{array}{|c|c|}\hline 1 & 1 \\ \hline\end{array} \quad \begin{array}{|c|c|c|}\hline 1 & 1 & 3 \\ \hline\end{array} \quad \begin{array}{|c|c|c|}\cline{1-1} 3 & \multicolumn{2}{c}{} \\ \hline 1 & 1 & 2 \\ \hline\end{array} \quad \begin{array}{|c|c|c|c|}\cline{1-1} 3 & \multicolumn{3}{c}{} \\ \hline 1 & 1 & 2 & 5 \\ \hline\end{array} \quad \begin{array}{|c|c|c|c|}\cline{1-2} 3 & 5 & \multicolumn{2}{c}{} \\ \hline 1 & 1 & 2 & 4 \\ \hline\end{array}$$

Hence $w_3 \neq_{P_5} w_1$, even though both words contain the same letters.

6.1.2 Introduction to plactic monoid research

In this chapter we will investigate the properties of finitely generated submonoids of the plactic monoid P_n. In particular we shall consider the following problems.

- Is every finitely generated submonoid of the plactic monoid finitely presented i.e. is the plactic monoid coherent?
- Does every finitely generated submonoid of the plactic monoid have FDT?
- Does every finitely generated submonoid of the plactic monoid admit a presentation by a finite complete presentation?

Recall that previously in Chapter 5 we saw that for the bicyclic monoid **B** the answer to all these questions is yes. In this chapter we obtain partial results to the three questions above by identifying certain families of submonoids of plactic monoids (for certain values of n) for which answers to the questions above are yes. One thing that makes this investigation more difficult than the corresponding result for the bicyclic monoid is that there are currently no results in the literature which classify all forms of submonoids of plactic monoids.

In [4] Cain, Gray and Malheiro proved that for all $n \geq 1$ the plactic monoid P_n

admits a finite complete presentation. The complete presentation they obtain is given with respect to a particular finite generating set called the column generators. Given this result, it is natural to ask whether the finitely generated submonoids of P_n also admit finite complete rewriting systems. In earlier work [30] Kubat and Okniński used the Knuth-Bendix completion procedure [3, Section 2.4] to prove that the plactic monoid P_3 admits a finite complete presentation with respect to the generating set $\{1, 2, 3\}$. They also prove that the same approach does not work for P_n where $n > 3$.

6.1.3 Existing results on the plactic monoid P_3

Theorem 6.1.6. *[30, Theorem 1 and Corollary 2] Let $B = \{1, 2, 3\}$ and let*

$$R = \{(332, 323), (322, 232), (331, 313), (311, 131), (221, 212), (211, 121), (231, 213),$$

$$(312, 132), (3212, 2321), (32131, 31321), (32321, 32132)\}.$$

Then (B, R) is a finite complete rewriting system defining the plactic monoid P_3. Moreover, with respect to this system, the irreducible words are precisely those of the form

$$(1)^i(21)^j(2)^k(321)^l(32)^m(3)^q \text{ or}$$

$$(1)^i(21)^j(31)^k(321)^l(32)^m(3)^q$$

for non-negative integers i, j, k, l, m, q.

6.2 New research on isomorphic submonoids of the plactic monoid

In this section we consider the plactic monoid P_3 and the submonoid generated by the set $X_q = \{1^q, 2^q, 3^q\}$ for some fixed $q \in \mathbb{N}$, and then go on to prove two new general results. First let us look at an example which will motivate the results that follow.

Example 6.2.1. Let P_3 be the plactic monoid on the alphabet $A = \{1, 2, 3\}$ with $w, w_3 \in A^+$ and $w \equiv 13221$, $w_3 \equiv 1^3 3^3 2^3 2^3 1^3$. Then we look at their tableau words as follows:

$$P(w) \equiv P(13221) \equiv 3\ 2\ 112 \quad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline\end{array} \quad \begin{array}{|c|c|}\cline{1-1} 3 \\ \hline 1 & 2 \\ \hline\end{array} \quad \begin{array}{|c|c|c|}\cline{1-1} 3 \\ \hline 1 & 2 & 2 \\ \hline\end{array} \quad \begin{array}{|c|c|c|}\cline{1-1} 3 \\ \cline{1-1} 2 \\ \hline 1 & 1 & 2 \\ \hline\end{array}$$

$$P(w_3) \equiv P(111333222222111) \equiv 333\ 222\ 111111222 \equiv 3^3\ 2^3\ 1^3 1^3 2^3$$

$$\begin{array}{|c|c|c|c|c|c|}\hline 1 & 1 & 1 & 3 & 3 & 3 \\ \hline\end{array} \quad \begin{array}{|c|c|c|c|c|c|}\cline{1-3} 3 & 3 & 3 \\ \hline 1 & 1 & 1 & 2 & 2 & 2 \\ \hline\end{array}$$

$$\begin{array}{|c|c|c|c|c|c|c|c|c|}\cline{1-3} 3 & 3 & 3 \\ \hline 1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 & 2 \\ \hline\end{array} \quad \begin{array}{|c|c|c|c|c|c|c|c|c|}\cline{1-3} 3 & 3 & 3 \\ \cline{1-3} 2 & 2 & 2 \\ \hline 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 \\ \hline\end{array}$$

So we start to see a pattern that suggests an isomorphism between the plactic monoid P_n and the submonoid generated by $\{1^q, 2^q, \ldots, n^q\}$ for a fixed $q \in \mathbb{N}$. First a useful lemma that proves how the q-blocks of numbers will always stay together when looking at the unique tableau form for an element.

Lemma 6.2.2. *Let P_n be the plactic monoid generated by the alphabet $A = \{1, 2, \ldots, n\}$ where $n \in \mathbb{N}$. Let $A_q = \{1^q, 2^q, \ldots, n^q\}$ with $q \in \mathbb{N}$. Let $g : A^* \to A^*$ be the homomorphism satisfying $g(a) = a^q$ for all $a \in A$. Then for every word $w \in A^*$, if w is a tableau word then $g(w)$ is also a tableau word.*

Proof. If w is a tableau word, it is the word of some tableau t. Let t^q be the tableau obtained by replacing each $a \in t$ with q occurrences of a. Then $g(w)$ is the word corresponding to t^q and hence is a tableau word. $\square$

Lemma 6.2.3. *Let P_n be the plactic monoid generated by the alphabet $A = \{1, 2, \ldots, n\}$ where $n \in \mathbb{N}$. Let S be the submonoid generated by the set $A_q = \{1^q, 2^q, \ldots, n^q\}$ and let $g : A^* \to A^*$ be the homomorphism satisfying*

$g(a) = a^q$ for all $a \in A$. Then, for every $w \in A^*$

$$P\big(g(w)\big) \equiv g\big(P(w)\big) \in A^*.$$

Proof. We proceed by induction on $|w|$.

Base case: Let $w \equiv a \in A$ so $|w| = 1$. Then

$$g\big(P(a)\big) \equiv g(a) \equiv a^q \text{ and } P\big(g(a)\big) \equiv P(a^q) \equiv a^q.$$

Inductive step: Write $w \equiv a_1a_2 \dots a_l\gamma \equiv w'\gamma$ with $\gamma, a_1, \dots, a_l \in A$ and $|w| \geq 2$, then $g(w) \equiv a_1^q \dots a_l^q\gamma^q$ and

$$\begin{aligned} P\big(g(w)\big) &\equiv P\big(\ g(w')\ g(\gamma)\ \big) && [\text{ since } g \text{ is a homomorphism }] \\ &\equiv P\big(\ P\big(g(w')\big)\ g(\gamma)\ \big) && [\text{ by definition of } P\ ] \\ &\equiv P\big(\ g\big(P(w')\big)\ g(\gamma)\ \big) && [\text{ by induction }] \\ &\equiv P\big(\ g\big(P(w')\ \gamma\big)\ \big) && [\text{ since } g \text{ is a homomorphism }]\ . \end{aligned}$$

Recall that we are aiming to show that $P\big(g(w)\big) \equiv g\big(P(w)\big)$ for every $w \in A^*$ and we have just shown above that $P\big(g(w)\big) \equiv P\big(g\big(P(w')\gamma\big)\big)$. So to complete the proof it suffices to prove that

$$P\Big(\ g\big(P(w')\gamma\big)\ \Big) \equiv g\big(\ P(w)\ \big).$$

Now by definition of P we know $P(w) \equiv P(w'\gamma) \equiv P\big(\ P(w')\ \gamma\ \big)$.

Set $\tau \equiv P(w')$ which is a tableau word by definition of P.

Now we need to prove that

$$P\big(\ g(\tau\gamma)\ \big) \equiv g\big(\ P(\tau\gamma)\ \big).$$

So to complete the proof of the lemma we shall now prove by induction on the number of rows that for every tableau word τ and every letter $\gamma \in A = \{1, \dots, n\}$

we have

$$P\big(\ g(\tau\gamma)\ \big) \equiv g\big(\ P(\tau\gamma)\ \big).$$

Base case: Suppose τ has just one row. Write $\tau \equiv a_1 a_2 \dots a_l \in A^*$ with $A = \{1, 2, \dots, n\}$. Since τ is a row word it satisfies $a_1 \leq a_2 \leq \dots \leq a_l$.

If $\gamma \geq a_l$ then $\tau\gamma$ is a tableau word and $g(\tau\gamma)$ is also a tableau word by Lemma 6.2.2 and consequently

$$g\big(P(\tau\gamma)\big) \equiv g(\tau\gamma) \equiv P\big(g(\tau\gamma)\big).$$

If $\gamma < a_l$ then with j minimal such that $a_j > \gamma$ then Schensted algorithm gives $P(\tau\gamma) \equiv a_j(a_1 \dots a_{j-1}\ \gamma\ a_{j+1} \dots a_l)$ so that

$$g\big(P(\tau\gamma)\big) \equiv a_j^q\ (a_1^q \dots a_{j-1}^q\ \gamma^q\ a_{j+1}^q \dots a_l^q).$$

On the other hand, we have $P\big(g(\tau\gamma)\big) \equiv P(a_1^q \dots a_l^q \gamma^q)$. Consider $P(a_1^q \dots a_l^q \gamma)$. Note that $g(\tau) \equiv a_1^q \dots a_l^q$ is a tableau word (Lemma 6.2.2). From the definition of j above, we see that $P\big(a_1^q \dots a_l^q \gamma\big) \equiv a_j\ \big(a_1^q \dots a_{j-1}^q\ (a_j^{q-1}\gamma)\ a_{j+1}^q \dots a_l^q\big)$. Repeating this argument, using Schensted's algorithm to insert each letter of γ^q, we eventually obtain

$$\begin{aligned} P\big(g(\tau\gamma)\big) &\equiv P(a_1^q \dots a_l^q \gamma^q) \\ &\equiv a_j^q(a_1^q \dots a_{j-1}^q\ \gamma^q\ a_{j+1}^q \dots a_l^q) \\ &\equiv g\big(P(\tau\gamma)\big), \text{ as required.} \end{aligned}$$

This completes the base case.

Inductive step: Let $\tau \equiv \alpha^{(1)} \dots \alpha^{(k)}$ be a tableau word with rows $\alpha^{(1)}, \dots, \alpha^{(k)}$ and $k \geq 2$, and let $\gamma \in A$. If $\alpha^{(k)}\gamma$ is a row word then $\tau\gamma$ is a tableau word, so

$P(\tau\gamma) \equiv \tau\gamma$, and hence $g(\tau\gamma)$ is also a tableau word (Lemma 6.2.2) and thus

$$\begin{aligned} P\big(g(\tau\gamma)\big) &\equiv g\big(\tau\gamma\big) \\ &\equiv g\big(P(\tau\gamma)\big), \text{ as required.} \end{aligned}$$

Now suppose that $\alpha^{(k)}\gamma$ is not a row. Write $\tau \equiv \tau'\alpha^{(k)}$ with $\alpha^{(k)} \equiv a_1 \ldots a_l$. So $\gamma < a_l$ with j minimal such that $a_j > \gamma$ then Schensted's algorithm gives $P(\alpha^{(k)}\gamma) \equiv a_j\ (a_1 \ldots a_{j-1}\ \gamma\ a_{j+1} \ldots a_l)$.

Now we have

$$\begin{aligned} g\big(P(\tau\gamma)\big) &\equiv g\big(P(\tau'\alpha^{(k)}\gamma)\big) \\ &\equiv g\big(P(\tau' a_j)\ \alpha'^{(k)}\big) \text{ where } \alpha'^{(k)} \equiv a_1 \ldots a_{j-1}\ \gamma\ a_{j+1} \ldots a_l \\ &\equiv g\big(P(\tau' a_j)\big)\ \ g(\alpha'^{(k)}) \qquad [\text{ since } g \text{ is a homomorphism }] \\ &\equiv P\big(g(\tau' a_j)\big)\ g(\alpha'^{(k)}) \qquad [\text{ by induction }] \\ &\equiv P\big(g(\tau')a_j^q\big)\ (a_1^q \ldots a_{j-1}^q\ \gamma^q\ a_{j+1}^q \ldots a_l^q) \\ &\equiv P\Big(g(\tau')\ P\big(g(\alpha^{(k)}\gamma)\big)\ \Big) \quad [\text{ by the base case }] \\ &\equiv P\big(\ g(\tau')\ g(\alpha^{(k)}\gamma)\ \big) \qquad [\text{ by definition of } P\ ] \\ &\equiv P\big(g(\tau\gamma)\big) \qquad\qquad [\text{ since } g \text{ is a homomorphism }]. \end{aligned}$$

This completes the proof. □

Theorem 6.2.4. *Let P_n be the plactic monoid generated by $A = \{1, 2, \ldots, n\}$ where $n \in \mathbb{N}$. Let S be the submonoid of P_n generated by $A_q = \{1^q, 2^q, \ldots, n^q\}$ for some fixed $q \in \mathbb{N}$. Then S is isomorphic to P_n.*

Proof. We have $A = \{1, 2, \ldots, n\}$ and $A_q = \{1^q, 2^q, \ldots, n^q\}$. As in Lemma 6.2.2 let $g : A^* \to A^*$ be the mapping defined by $g(a) = a^q$ for all $a \in \{1, 2, \ldots, n\}$. We claim that g induces a well-defined isomorphism between P_n and S. By Lemma 6.2.3 we know that for all words $w \in A^*$ we have $P\big(g(w)\big) \equiv g\big(P(w)\big)$.

Let $w_1, w_2 \in A^*$. If $w_1 = w_2$ in P_n then $P(w_1) \equiv P(w_2)$ which implies

$g\big(P(w_1)\big) \equiv g\big(P(w_2)\big)$ and hence

$P\big(g(w_1)\big) \equiv g\big(P(w_1)\big) \equiv g\big(P(w_2)\big) \equiv P\big(g(w_2)\big)$ which implies $g(w_1) = g(w_2)$ in P_n. This proves that g induces a well-defined homomorphism $\hat{g}$ from P_n to S. Since S is generated by $g(\{1, \ldots, n\})$ the map $\hat{g}$ is surjective. Finally, to see that $\hat{g}$ is injective, if $g(w_1) = g(w_2)$ in P_n then

$$
\begin{aligned}
& P\big(g(w_1)\big) \equiv P\big(g(w_2)\big) \\
\Rightarrow\ & g\big(P(w_1)\big) \equiv g\big(P(w_2)\big) \\
\Rightarrow\ & P(w_1) \equiv P(w_2) \quad [\text{ since } g : A^* \to A^* \text{ is injective }] \\
\Rightarrow\ & w_1 = w_2 \text{ in } P_n.
\end{aligned}
$$

This completes the proof that $\hat{g} : P_n \to S$ is an isomorphism. $\square$

We now aim to generalise Theorem 6.2.4. Let us begin with an illustrative example.

Example 6.2.5. Let P_3 be the plactic monoid on the alphabet $A = \{1, 2, 3\}$ and $w_1 \equiv 132$, $w_2 \equiv 13^3 2^2$, $w_3 \equiv 1^5 3^{15} 2^{10}$ with $w_1, w_2, w_3 \in A^+$. Then we look at their tableau forms as follows:

$$P(w_1) \equiv P(132) \equiv 3\ 12 \quad \begin{array}{|c|}\hline 1 \\ \hline\end{array} \quad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline\end{array} \quad \begin{array}{|c|c}\hline 3 & \multicolumn{1}{|c}{} \\ \hline 1 & \multicolumn{1}{|c|}{2} \\ \hline\end{array}$$

$$P(w_2) \equiv P(13^3 2^2) \equiv 33\ 1223 \quad \begin{array}{|c|}\hline 1 \\ \hline\end{array} \quad \begin{array}{|c|c|c|c|}\hline 1 & 3 & 3 & 3 \\ \hline\end{array} \quad \begin{array}{|c|cccc}\cline{1-1} 3 & & & & \\ \hline 1 & \multicolumn{1}{|c|}{2} & \multicolumn{1}{c|}{3} & \multicolumn{1}{c|}{3} \\ \hline\end{array} \quad \begin{array}{|c|c|cc}\cline{1-2} 3 & 3 & & \\ \hline 1 & 2 & \multicolumn{1}{|c|}{2} & \multicolumn{1}{c|}{3} \\ \hline\end{array}$$

$$P(w_3) \equiv P(1^5 3^{15} 2^{10}) \equiv 3^5 3^5\ 1^5 2^5 2^5 3^5$$

$$\begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 1&1&1&1&1&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3 \\ \hline\end{array}$$

$$\begin{array}{|c|ccccccccccccccccccc}\cline{1-1} 3 &&&&&&&&&&&&&&&&&&& \\ \hline 1&\multicolumn{1}{|c|}{1}&\multicolumn{1}{c|}{1}&\multicolumn{1}{c|}{1}&\multicolumn{1}{c|}{1}&\multicolumn{1}{c|}{2}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3} \\ \hline\end{array}$$

$$\begin{array}{|c|c|c|c|c|ccccccccccccccc}\cline{1-5} 3&3&3&3&3&&&&&&&&&&&&&&& \\ \hline 1&1&1&1&1&\multicolumn{1}{|c|}{2}&\multicolumn{1}{c|}{2}&\multicolumn{1}{c|}{2}&\multicolumn{1}{c|}{2}&\multicolumn{1}{c|}{2}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3}&\multicolumn{1}{c|}{3} \\ \hline\end{array}$$

$$\begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}
\cline{1-6}
3&3&3&3&3&3 \\
\hline
1&1&1&1&1&2&2&2&2&2&2&3&3&3&3&3&3&3&3&3 \\
\hline
\end{array}$$

$$\begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}
\cline{1-10}
3&3&3&3&3&3&3&3&3&3 \\
\hline
1&1&1&1&1&2&2&2&2&2&2&2&2&2&2&3&3&3&3&3 \\
\hline
\end{array}$$

So we start to see a pattern that suggests an isomorphism between the submonoids of the plactic monoid P_3 generated by the sets $A_2 = \{1, 2^2, 3^3\}$ and $A_3 = \{1^5, 2^{10}, 3^{15}\}$. Note, it can clearly be seen that there is no isomorphism between P_3 and either the submonoid generated by A_2, or the submonoid generated by A_3, as there is no common divisor across the generating alphabet A_2 (or A_3) when compared to A. Also, we observe that the generators $3^3 \in A_2$ and $3^{15} \in A_3$ have been "split" when we look at the tableau word, meaning that in general, if $b_2 \in A_2^*$ and $b_3 \in A_3^*$ then we can have $P(b_2) \notin A_2^*$ and $P(b_3) \notin A_3^*$. However, in tableau word form the individual letters from the alphabet A remain in blocks of five, which is a common divisor of the powers in the set A_3. In addition, if we divide the powers of all the generators in A_3 by five we get the generating set A_2. This is similar to what we have seen previously in this section.

Theorem 6.2.6. *Let P_n be the plactic monoid generated by $A = \{1, 2, \ldots, n\}$ where $n \in \mathbb{N}$. Let S be the submonoid of P_n generated by $A_s = \{1^{s_1}, 2^{s_2}, \ldots, n^{s_n}\}$ for $s_1, s_2, \ldots, s_n \in \mathbb{N}$ and set $q = gcd\{s_1, s_2, \ldots, s_n\}$. Let T be the submonoid generated by $A_t = \{1^{t_1}, 2^{t_2}, \ldots, n^{t_n}\}$ where $t_x = s_x/q$ for all $x \in \{1, 2, \ldots, n\}$.*

Then S is isomorphic to T.

Proof. This is proved in a similar way to Theorem 6.2.4. We shall just highlight the changes needed to modify that proof.

In order to make this proof clearer to see we redefine the generating sets for the submonoids S and T as follows:

Let T be generated by the set $B = \{1^{t_1}, 2^{t_2}, \ldots, n^{t_n}\}$.

Let S be generated by the set $B_q = \{(1^q)^{t_1}, (2^q)^{t_2}, \ldots, (n^q)^{t_n}\}$.

Let $g : A^* \to A^*$ be the homomorphism defined by $g(a) = a^q$ for all $a \in A$. From Lemma 6.2.3 it follows that $g(P(w)) \equiv P(g(w))$ for all $w \in A^*$. Observe that g maps B to B_q and hence maps B^* to B_q^*. Using the fact that $g(P(w)) \equiv P(g(w))$ for all $w \in B^*$ it can then be shown (in a similar way to the proof of Theorem 6.2.4) that g induces an isomorphism between T and S. $\square$

6.3 Monoids with homogeneous presentations

6.3.1 Introduction

The results in this section arise from investigating the properties of a different submonoid of the plactic monoid P_n, that is the submonoid which consists of all the elements of P_n which have a length divisible by j where $j \in \mathbb{N}$ is fixed. It is clear from the definition of the plactic monoid that all words that represent the same element will have the same length, therefore we can refer to the *length of an element* and *element length*. So we consider the submonoid generated by the set $A_j = \{a_1 a_2 \dots a_j : a_1, a_2, \dots, a_j \in A\}$.

This motivates a proposed general theory regarding a particular form of presentation for monoids, which we define below.

Definition 6.3.1. Let $\langle A \mid R \rangle$ be a presentation where A is an alphabet and

$$R = \{(u_1, v_1), (u_2, v_2), \dots, (u_k, v_k)\} \text{ and } |u_i| = |v_i| \text{ for all } i \in \{1, 2, \dots, k\}.$$

Then $\langle A \mid R \rangle$ is a *homogeneous presentation.*

In this section we will prove the following result.

Proposition 6.3.2. *Let M be the monoid defined by the homogeneous presentation $\langle A \mid R \rangle$ where A is a finite alphabet and R is a finite set of relations in $A^* \times A^*$. Fix $j \in \mathbb{N}$ and let E be the submonoid of M which consists of all the elements of M which have a length divisible by j.*

Suppose further that the presentation $\langle A \mid R \rangle$ is a finite complete rewriting system. Then there exists a presentation for E which is a finite complete rewriting system.

Recall that by Theorem 6.1.6 the plactic monoid P_3 admits a finite homogeneous presentation which is a finite complete rewriting system. As such we will be able to apply the above proposition, in the relevant circumstances, once we have proved it is true.

Throughout the remainder of Section 6.3 the definitions in Proposition 6.3.2 for the monoid M and submonoid E will apply unless stated otherwise. In addition, let η be the smallest congruence on A^* which contains R. Thus if $w \in A^*$ is a word, then w/η is the element which is represented by w in M.

An obvious consequence of a homogeneous presentation is that all the words which represent the same element will have the same length as the relations do not alter the length of the word. (The reverse is not true i.e. all words of the same length do not necessarily represent the same element.) We can refer to *the length of an element* of a monoid which is defined by a homogeneous presentation by defining it to be the length of a word by which it is represented. This notion of *element length* is well defined as it is independent of the choice of word. Next, a useful lemma, the proof for which is obvious.

Lemma 6.3.3. *Let M be a monoid defined by a finite homogeneous presentation and E the submonoid consisting of all the elements of M whose length is divisible by some fixed $j \in \mathbb{N}$. Define the set of words in A^* that are of length j:*

$$X_0 = \{ \, w \in A^* : |w| = j \, \}.$$

Then X_0 is a finite generating set for E.

6.3.2 Linking new research to that in Chapter 4

Next we look at some useful properties of the monoid M and its submonoid E which links our questions to the results contained in Chapter 4. Recall the definitions from Chapter 4 regarding left (respectively right) unitary subsemigroups with finite strict right (respectively left) boundary.

We have been considering a proposition for a monoid with homogeneous presentation $\langle A \mid R \rangle$ and our previous definitions and proofs in Chapter 4 are for a semigroup. We shall now apply the results from Chapter 4 to deduce some results about semigroups defined by homogeneous semigroup presentations and their subsemigroups. Then we will explain how the corresponding results for monoid presentations can be deduced.

Lemma 6.3.4. *Let $\langle B \mid Q \rangle$ be a finite homogeneous semigroup presentation which defines the semigroup S. Fix $j \in \mathbb{N}$ and let F be the subsemigroup of S which consists of all elements of S with length divisible by j.*

Then F is a left (respectively right) unitary subsemigroup of S with finite strict right (respectively left) boundary in S with respect to B.

Proof. Let $w \in \mathcal{L}(B, F)$ be an arbitrary word such that $|w| = mj$ where $m \in \mathbb{N}$. Let $w \equiv w_1 w_2$ where w_1 is the shortest prefix of w such that $w_1 \in \mathcal{L}(B, F)$. Then by definition of F we have $|w_1| = j$. Which means that $|w_2| = (m-1)j$ and therefore $w_2 \in \mathcal{L}(B, F)$. Thus F is a left unitary subsemigroup of S.

Let ρ be the smallest congruence on B^+ which contains Q. Let $u \in \mathcal{L}(B, F)$ be an arbitrary word such that $u \in \mathcal{SWB}_r(B, F)$. Then by definition, no proper prefix of u belongs to $\mathcal{L}(B, F)$. This implies that $|u| = j$ and therefore $\mathcal{SWB}_r(B, F)$ is a finite set. Thus, as $\mathcal{SB}_r(B, F) = \mathcal{SWB}_r(B, F)/\rho$, we can see that the strict right boundary of F in S is finite.

Therefore F is a left unitary subsemigroup of S with finite strict right boundary in S with respect to B. By a dual proof we can then see that F is a right unitary subsemigroup of S with finite strict left boundary in S with respect to B. □

Now that we have established this property for the subsemigroup F, we can apply results from Chapter 4 as follows in the next two corollaries.

Corollary 6.3.5. *Let $\langle B \mid Q \rangle$ be a finite homogeneous semigroup presentation which defines the semigroup S. Fix $j \in \mathbb{N}$ and let F be the subsemigroup of S which consists of all elements of S with length divisible by j.*

Then F is finitely presented.

Proof. By Lemma 6.3.4 and Theorem 4.5.1. □

Corollary 6.3.6. *Let $\langle B \mid Q \rangle$ be a finite homogeneous semigroup presentation which defines the semigroup S. Fix $j \in \mathbb{N}$ and let F be the subsemigroup of S which consists of all elements of S with length divisible by j.*

If S has finite derivation type then F also has finite derivation type.

Proof. By Lemma 6.3.4 and Theorem 4.1.1. □

Since in this section we are interested in monoids defined by homogeneous monoid presentations (as plactic monoids are defined this way), we now observe that all the results in this Subsection 6.3.2 also hold for homogeneous monoid presentations. First we include some existing results which will be used in the proof of the results for homogeneous monoid presentations.

Theorem 6.3.7. *[53, Theorem 2] Let S be a small extension of T. If T has finite derivation type (FDT), then so does S.*

Recall that if T is a *large subsemigroup* of the semigroup S then $S \setminus T$ is finite. If T is an ideal of S which is also a large subsemigroup of S, then T is called a *large ideal* of S.

Theorem 6.3.8. *[40, Theorem 1] Let S be a semigroup and let T be a large ideal of S. If S has FDT, then so does T.*

Theorem 6.3.9. *[4, Theorem 3.4] There exists a finite complete rewriting system which defines the plactic monoid P_n.*

Corollary 6.3.10. *[4, Corollary 3.5] Every plactic monoid has finite derivation type.*

Now we prove two new results which are for homogeneous monoid presentations and are the monoid analogue of Corollaries 6.3.5 and 6.3.6.

Corollary 6.3.11. *Let $\langle A \mid R\rangle$ be a finite homogeneous monoid presentation defining a monoid M. Fix $j \in \mathbb{N}$ and let E be the submonoid of M which consists of all elements of M with length divisible by j. Then:*
(i) E is a finitely presented monoid and
(ii) if M has finite derivation type then E has finite derivation type.

Proof. Let S be the semigroup defined by $\langle A \mid R\rangle$, regarded as a semigroup presentation. Since $\langle A \mid R\rangle$ is homogeneous it follows that the only word in A^* that represents the identity element of M is the empty word ϵ (since no defining relations from R can be applied to the empty word ϵ). Note that it makes sense to view $\langle A \mid R\rangle$ as a semigroup presentation since being a homogeneous presentation implies $(u, v) \in A^+ \times A^+$ for all relations $(u, v) \in R$. It follows that for any two words $\alpha, \beta \in A^+$ we have $\alpha = \beta$ in M if and only if $\alpha = \beta$ in S. It follows that S is isomorphic to the subsemigroup of M generated by all the non-identity elements of M. So $S \cong M \setminus \{1\}$ where 1 is the identity of M. So S is isomorphic to a subsemigroup of M with finite complement and by Lemma 2.5.1 there exists a semigroup presentation for the monoid M. But then it follows by Theorem 2.6.18 that S is a finitely presented semigroup if and only if M is a finitely presented monoid. It can easily be proved that S is a large ideal of M and so by Theorem 6.3.8 if M has FDT then S has FDT.

Now let F be the subsemigroup of M generated by all elements of M with length divisible by j. Then F is the subsemigroup of E given by all the non-empty words, that is, $F \cong E \setminus \{1\}$. By Lemma 2.5.1 there exists a semigroup presentation for the monoid E. Hence applying Theorem 2.6.18, E is a finitely presented monoid if and only if F is a finitely presented semigroup, and by Lemma 6.3.7 the monoid E has FDT if the semigroup F has FDT. Combining these observations with Corollaries 6.3.5 and 6.3.6 we conclude:

(i) F is a finitely presented semigroup (by Corollary 6.3.5), since S is a finitely presented semigroup and F is isomorphic to the subsemigroup of S of all elements of length divisible by j, and hence (by Theorem 2.6.18 and Lemma 2.5.1) E is a finitely presented monoid; and

(ii) if M has FDT then the semigroup S has FDT (by Theorem 6.3.8) which implies the semigroup F has FDT (by Crollary 6.3.6) which implies that the monoid E has FDT (by Lemma 6.3.7). $\square$

Now we have a result which we can apply to the plactic monoid as it has a homogeneous presentation. We can easily apply this result to the plactic monoid P_3 but in this case we have a result which can be extended to P_n.

Corollary 6.3.12. *Let P_n be the plactic monoid of rank n. Fix $j \in \mathbb{N}$ and let E be the submonoid of P_n which consists of all the elements of P_n with length divisible by j. Then:*
(i) E is a finitely presented monoid, and
(ii) the monoid E has finite derivation type.

Proof. Part (i) follows from Corollary 6.3.11.
For part (ii), by Theorem 6.3.9, the plactic monoid P_n admits a FCRS and hence (by Corollary 6.3.10) has FDT. Now (ii) follows from Corollary 6.3.11 (ii). $\square$

In the next section we move on to investigate whether this corollary also holds for the property of admitting a finite complete rewriting system (FCRS).

6.3.3 Codified submonoids

This section provides an overview of relevant results from the paper [37]. The results will provide a process by which we can prove Proposition 6.3.2. This proposition could also be proved using Reidemeister-Schreier type rewriting methods (see Theorem 2.9.1), but applying the results from [37] gives a shorter, more elegant proof. We retain the notation from earlier in this section, so $\langle A \mid R \rangle$ is a finite homogeneous monoid presentation defining a monoid $M \cong A^*/\eta$ where η is the congruence of A^* generated by R, and E is the submonoid of M which consists of all elements of M with length divisible by j, for some fixed $j \in \mathbb{N}$. We first introduce some new definitions and then the relevant result, all of which come from [37].

A subset X_0 of A^+ is said to be a *code* in the alphabet A if the submonoid of A^* generated by X_0, $\langle X_0 \rangle$, is free on X_0. Therefore, if X_0 is a code in the alphabet A, we have $\langle X_0 \rangle \cong X_0^*$. The idea behind this concept is quite simple, we want to think of the elements of X_0 as being codified by words of the alphabet A. Hence, the set X_0 is a code if and only if there is no word in $\langle X_0 \rangle$ that can be decoded in two different ways as elements of X_0, see the following Example 6.3.13.

Example 6.3.13. Let $A = \{a, b, c\}$ be an alphabet and define the two generating sets $X_1 = \{aa, ab, ba, ac, ca\}$ and $X_2 = \{a, ab, bc, c\}$.

If we take a word in $\langle X_1 \rangle$, say $baaaca$, we can see that X_1 is a code as the word can only be decoded in one way as an element in $\langle X_1 \rangle$, that is by reading two letters at a time. Whether we read from the left to right or from right to left, we will identify the same elements from X_1. However, if we consider the words in $\langle X_2 \rangle$ we can easily come up with an example that shows it is not a code. The word abc can be decoded in two ways as $(ab)\ c$ or $a\ (bc)$, which clearly identifies different elements in $\langle X_2 \rangle$.

The set of the left terms of R, that is, $\{r^{+1} \in A^* \mid \exists r^{-1} \in A^* : (r^{+1}, r^{-1}) \in R\}$ is denoted by $dom(R)$. By $im(R)$ we denote the set of right terms of R, that is,

$\{r^{-1} \in A^* \mid \exists r^{+1} \in A^* : (r^{+1}, r^{-1}) \in R\}$.

Let w be a word from A^+ and let X_0 be a subset of A^+. Then we say that w is *extensible* to X_0 if w is a factor of some word in $\langle X_0 \rangle$, that is, if there exists words $x, y \in A^*$ such that $xwy \in \langle X_0 \rangle$.

Let w be a word that is extensible to X_0. We say that a word $u \in \langle X_0 \rangle$ is a *minimal extension* of w in X_0 if (i) w is a factor of u, and (ii) if $u \equiv x_1 x_2 w y_1 y_2$ with $x_1,\ x_2 w y_1,\ y_2 \in \langle X_0 \rangle$ then $x_1 \equiv y_2 \equiv 1$, where 1 is the empty word.

Observe that an extensible word has always a minimal extension. Actually it may have more than one minimal extension. Clearly, if w is a word in $\langle X_0 \rangle$ then its unique minimal extension is w.

Next we construct a new set of rewrite rules R_0 from R. Let $r = (r^{+1}, r^{-1}) \in R$ be such that r^{+1} is extensible to $\langle X_0 \rangle$. For each minimal extension s of r^{+1} in X_0 we will define a new rewriting rule r_s in the following way: let $x, y \in A^*$ be such that $s \equiv x r^{+1} y$; we choose r_s as being the rewriting rule (s^{+1}, s^{-1}) where $s^{+1} \equiv s$ and s^{-1} is a word such that $x r^{-1} y \xrightarrow{*}_R s^{-1}$. Note that in general $x r^{-1} y \notin \langle X_0 \rangle$ so we look to find s^{-1} where $s^{-1} \in \langle X_0 \rangle$, see [37, Page 211 Examples]. Let R_0 be the set of all rewriting rules constructed in the previous way. Notice that $dom(R_0) \subseteq \langle X_0 \rangle$ and that the congruence $\xleftrightarrow{*}_{R_0}$ is contained in $\xleftrightarrow{*}_R$ since $\xrightarrow{*}_{R_0} \subseteq \xrightarrow{*}_R$.

On the sets X_0 and R_0, we consider the following conditions:

(A) X_0 is a code;

(B) $im(R_0) \subseteq \langle X_0 \rangle$.

Let us consider a new alphabet $\tilde{X}_0 = \{\tilde{x} : x \in X_0\}$ in one-to-one correspondence with X_0. Assuming condition (A), it is possible to extend this correspondence to an isomorphism between the monoids $\langle X_0 \rangle$ and $\tilde{X}_0^*$. We denote by $\tilde{u}$ an element $\tilde{x}_1 \dots \tilde{x}_m$ of $\tilde{X}_0^*$, where $m \in \mathbb{N}$ and $x_i \in X_0$. Notice that we have $\tilde{u} \equiv \widetilde{x_1 \dots x_m} \equiv \tilde{x}_1 \dots \tilde{x}_m$, where $u \equiv x_1 \dots x_m$.

Assuming (A) and (B), to each pair $(r^{+1}, r^{-1}) \in R_0$, there corresponds a unique pair $(\tilde{r}^{+1}, \tilde{r}^{-1}) \in \tilde{X}_0^* \times \tilde{X}_0^*$. Note that the uniqueness here follows from assumption (A) that X_0 is a code. Hence, we can define a new set of rewrite rules

$$\tilde{R}_0 = \{ \ (\tilde{r}^{+1}, \tilde{r}^{-1}) \mid (r^{+1}, r^{-1}) \in R_0 \ \}.$$

Notice that if R_0 is finite, so is $\tilde{R}_0$. Now we have enough definitions from [37] and can state the main theorem which follows:

Theorem 6.3.14. *[37, Theorem 2.4] Let (A, R) be a complete rewriting system defining the monoid M. Let $X_0 \subseteq A^+$ be such that conditions (A) and (B) hold. Then the rewriting system $(\tilde{X}_0, \tilde{R}_0)$ is complete and defines the submonoid of M generated by X_0.*

Note that it is not a condition of the above theorem that $dom(R)$ is extensible, see [37, Page 211 Examples] for several examples.

6.3.4 New research relating to finite homogeneous presentations

In this section we return to our Proposition 6.3.2 and prove it as the following new theorem.

Theorem 6.3.15. *Let M be the monoid defined by the homogeneous presentation $\langle A \mid R \rangle$ where A is a finite alphabet and R is a finite set of relations in $A^* \times A^*$. Fix $j \in \mathbb{N}$ and let E be the submonoid of M which consists of all elements of M which have a length divisible by j.*

Suppose further that the presentation $\langle A \mid R \rangle$ is a finite complete rewriting system. Then there exists a presentation for E which is a finite complete rewriting system.

Proof. The proof appeals to Theorem 6.3.14 and as such it suffices to prove that conditions (A) and (B) hold and that X_0 and R_0 are finite. Let $X_0 = \{w \in A^* : |w| = j\}$. Then by Lemma 6.3.3 $\langle X_0 \rangle$ generates E.

(A) X_0 is a code

Let $w \in A^*$. If $w \in \langle X_0 \rangle$ then $|w| = mj$ for some $m \in \mathbb{N}$ with $m \geq 1$. Write $w \equiv \alpha_1 \alpha_2 \ldots \alpha_m$ where $\alpha_i \in X_0$ for all $i \in \{1, 2, \ldots, m\}$. If $\beta_1 \beta_2 \ldots \beta_l \equiv w$ with $\beta_i \in X_0$ for $i \in \{1, 2, \ldots, l\}$, then $\beta_1 \beta_2 \ldots \beta_l \equiv \alpha_1 \alpha_2 \ldots \alpha_m$. By definition of X_0 we have $|\beta_i| = j = |\alpha_k|$ for all $i \in \{1, 2, \ldots, l\}$ and $k \in \{1, 2, \ldots, m\}$. It follows that $m = l$ and $\alpha_i = \beta_i$ for all $i \in \{1, 2, \ldots, m\}$. Hence $w \equiv \alpha_1 \alpha_2 \ldots \alpha_m$ is the unique way of writing $w \in A^*$ as a product of words from X_0. Hence X_0 is a code.

(B) $im(R_0) \subseteq \langle X_0 \rangle$

The rewrite rules R for the monoid M are homogeneous and therefore for all $(r^{+1}, r^{-1}) \in R$ we have $|r^{+1}| = |r^{-1}|$. Let $(r^{+1}, r^{-1}) \in R$. For every minimal extension s of r^{+1} write $s \equiv xr^{+1}y$ in A^* and define r_s to be the rewriting rule (s^{+1}, s^{-1}) where $s^{+1} \equiv xr^{+1}y$ and $s^{-1} \equiv xr^{-1}y$. Note that trivially we have $xr^{-1}y \xrightarrow{*}_R s^{-1}$. Hence R_0 can be chosen to be equal to the set of all these relations r_s. Since R is homogeneous $|xr^{+1}y| = |xr^{-1}y|$ which is a multiple of j since $xr^{+1}y \in X_0^*$. It follows that $s^{-1} \equiv xr^{-1}y \in \langle X_0 \rangle$, for every r_s in R_0. Hence condition (B) holds.

Fact: Every word $w \in A^*$ has only finitely many minimal extensions to X_0.

Proof of fact: Suppose $u \in \langle X_0 \rangle$ is a minimal extension of w to X_0. Write $u \equiv \beta_1 \beta_2 \ldots \beta_m$ where $|\beta_i| = j$ for all $i \in \{1, 2, \ldots, m\}$. Since this extension is minimal it follows that w is a subword of u where w contains $\beta_2 \beta_3 \ldots \beta_{m-1}$ as a strict subword. It follows that $m \leq |w|$. Since $|w|$ and j are fixed and A is finite, it follows that there are only finitely many possible choices for u, completing the proof.

Corollary: Since R is finite by assumption, applying the above fact it follows that R_0 is finite.

It remains to observe that the sets A, R, X_0 are finite and R_0 is finite by the above Corollary and therefore so are $\tilde{X}_0$ and $\tilde{R}_0$. This completes the proof. $\square$

Corollary 6.3.16. *Let P_3 be the plactic monoid of rank 3 defined by the ordered alphabet $A = \{1, 2, 3\}$ and let*

$$R = \{(332, 323), (322, 232), (331, 313), (311, 131), (221, 212), (211, 121),$$

$$(231, 213), (312, 132), (3212, 2321), (32131, 31321), (32321, 32132)\}.$$

Then P_3 is defined by the presentation $\langle A \mid R \rangle$. Fix $j \in \mathbb{N}$ and let E_j be the submonoid of P_3 containing only elements of P_3 of length divisible by j.

Then there exists a presentation for E_j which is a finite complete rewriting system.

Proof. By Theorem 6.1.6 the presentation $\langle A \mid R \rangle$ defines the plactic monoid P_3 and is a finite complete rewriting system.

Note that this presentation is homogeneous and so by Theorem 6.3.15 there exists a presentation for E_j which is a finite complete rewriting system. □

Remark: In Corollary 6.3.16 above we have defined the monoid E_j by saying that it contains all the elements of P_3 of length divisible by j. To be in line with the other submonoids that we consider in this chapter, the monoid E_j can equally be defined as the submonoid of P_3 generated by the set of all tableau with exactly j boxes. To see this is true, refer to Lemma 6.3.3.

6.3.5 Further applications

Natural numbers under addition

Theorem 6.3.17. *Let M be a monoid defined by a finite homogeneous presentation $\langle A \mid R \rangle$. Let $\psi : M \to (\mathbb{N}, +)$ be the surjective homomorphism induced by the mapping $w \mapsto |w|$ for $w \in A^*$. Let T be a finitely generated submonoid of $(\mathbb{N}, +)$. Then* *(i)* $N = T\psi^{-1}$ *is a finitely presented submonoid of* M*.*

(ii) Moreover, if (A, R) *is a finite complete rewriting system then* $N = T\psi^{-1}$ *admits a presentation by a finite complete rewriting system.*

Proof. We prove both parts together:

Part (i) and (ii)

We begin with a useful quote from a recent paper [9]:
"The subsemigroups of the free monogenic semigroup $\mathbb{N} = \{1, 2, 3, \ldots\}$ are also reasonably tame, even though they are more complicated than subgroups of $\mathbb{Z}$. Every such subsemigroup has the form $A \cup B$, where A is finite, and $B = \{nd : n \geq n_0\}$ for some $n_0, d \in \mathbb{N}$; see [50]."

Define the set $C = \{c \in T : c < \kappa\}$ and $D = \{d \in T : d \geq \kappa\}$ such that $|C|$ is finite but $|D|$ is infinite and $T = C \cup D$. (In terms of the quoted result above, the set C corresponds to the set A and the set D to the set B. We use f instead of d, where f is the greatest common divisor of the numbers in the finite generating set for T. We set $\kappa = n_0 f$ where n_0 is fixed as the smallest integer such that for any $x, y \in T$ where $x, y \geq n_0 f$, then $|x - y| \equiv 0 \pmod{f}$. Thus κ is the number above which we see a predictable pattern for those numbers in the set T i.e. the numbers in T that are higher than κ are all multiples of f.) Let the elements in D be defined in ascending order i.e. $D = \{d_1, d_2, d_3 \ldots\}$. Then D is a subsemigroup of T with the finite generating set $X_D = \{d_1, d_2, \ldots, d_{2d_1-1}\}$.

Now consider $T\psi^{-1} = C\psi^{-1} \cup D\psi^{-1}$ and note that $T\psi^{-1} \setminus D\psi^{-1}$ is finite. Recall that $D\psi^{-1}$ comprises of all the words in A^* that are of length l where we have $l \geq \kappa$ and $l \equiv 0 \pmod{f}$. Define F to be a submonoid of M which comprises of all the elements of length that is divisible by f.

Given that M is defined by a finite homogeneous presentation then by Corollary 6.3.11 part (i) we can say that F is defined by a finite presentation. Given that M admits a presentation that is a FCRS then by Theorem 6.3.15 we know that

F also admits a presentation that is a FCRS.

Next we look at $D\psi^{-1}$ which can be seen is a subsemigroup of F where $|F \setminus D\psi^{-1}|$ is finite. By Theorem 2.6.18, as F is finitely presented, then $D\psi^{-1}$ is finitely presented. Also, by Theorem 2.6.20, as F admits a presentation which is a FCRS then $D\psi^{-1}$ also admits a presentation which is a FCRS.

Finally, we have $T\psi^{-1} = C\psi^{-1} \cup D\psi^{-1}$ with $D\psi^{-1}$ a subsemigroup of $T\psi^{-1}$ and $|T\psi^{-1} \setminus D\psi^{-1}|$ being finite. So by Theorem 2.6.18, as $D\psi^{-1}$ is finitely presented then $T\psi^{-1}$ is finitely presented. Also, by Theorem 2.6.19, as $D\psi^{-1}$ admits a presentation that is a FCRS, then $T\psi^{-1}$ also admits a presentation that is a FCRS. □

Corollary 6.3.18. *Let* $A = \{1, 2, 3\}$ *be an ordered alphabet and*

$$R = \{(332, 323), (322, 232), (331, 313), (311, 131), (221, 212), (211, 121), (231, 213),$$

$$(312, 132), (3212, 2321), (32131, 31321), (32321, 32132)\}.$$

Then the plactic monoid P_3 *of rank 3 is defined by the homogeneous presentation* $\langle A \mid R \rangle$.

Let $\psi : P_3 \to (\mathbb{N}, +)$ *be the surjective homomorphism induced by the mapping* $w \mapsto |w|$ *for* $w \in A^*$. *Let* T *be a finitely generated submonoid of* $(\mathbb{N}, +)$ *and let* $N = T\psi^{-1}$ *be a submonoid of* P_3. *Then* N *admits a presentation by a finite complete rewriting system.*

Proof. By Theorem 6.1.6 P_3 is defined by the finite complete rewriting system (A, R). Theorem 6.3.17 part (ii) completes the proof. □

6.4 Introduction to new research relating to submonoids of the plactic monoid P_2

We now expand our investigation to consider other families of finitely generated submonoids of the plactic monoid not already covered by earlier results in this chapter. We will focus our attention on submonoids generated by powers of the generators $\{1, 2, \ldots, n\}$, that is, generating sets of the form

$$X = \{1^{i_1}, 1^{i_2}, \ldots, 1^{i_r}, 2^{j_1}, 2^{j_2}, \ldots, 2^{j_s}, 3^{k_1}, 3^{k_2}, \ldots, 3^{k_t}\}.$$

This class of submonoids already exhibit complex behaviour and are difficult to study in general. Because of this we will restrict our attention to submonoids of the plactic monoid P_2. The submonoids considered are generated by various finite sets, working towards generalising any results to generating sets of the form $Y = \{1^i, 2^j\}$ where $i, j \in \mathbb{N}$ such that $i \leq j$.

We will prove that a finite presentation exists for specific submonoids of this form where the presentation is a FCRS. This naturally gives the result that it has a finite presentation and also FDT (by Theorem 4.4.2).

Research in the remainder of this chapter makes use of a result from the paper [30] which we referenced earlier in Theorem 6.1.6. This theorem gives a presentation for the plactic monoid P_3 which is a finite complete rewriting system. Next we include a corollary which follows from this theorem by restricting the presentation to the generators $\{1, 2\}$ and the relations to those that involve only 1 and 2.

Corollary 6.4.1. *[30, Theorem 1 and Corollary 2]*

Let $B = \{1, 2\}$ *and* $R = \{(221, 212), (211, 121)\}$.
Then (B, R) *is a finite complete rewriting system defining the plactic monoid* P_2. *Moreover, with respect to this system, the irreducible words are precisely those of the form* $(1)^i(21)^j(2)^k$ *for non-negative integers* i, j, k.

6.5 Submonoid of the plactic monoid P_2 generated by the set $X = \{1, 2^i\}$

6.5.1 Statement of new proposition

Proposition 6.5.1. *Let $A = \{1, 2\}$ and $R = \{(221, 212), (211, 121)\}$. Then the plactic monoid P_2 is defined by the presentation $\langle A \mid R \rangle$.*

Fix $i \in \mathbb{N}$ with $i > 1$ and let E be the submonoid of P_2 generated by $X = \{1, 2^i\}$. Set $B = \{a, b\}$ and let Q be the subset of $B^* \times B^*$ consisting of all the following pairs:
(i) (bba, bab),
(ii) (ba^{i+1}, aba^i),
(iii) $\{(ba^{j-1}ba, ba^j b) \ : \ 2 \leq j \leq i\}$.
Then (B, Q) is a finite complete rewriting system defining the monoid E where a and b correspond to the generators 1 and 2^i, respectively.

6.5.2 Outline proof and definitions

At this point it is worth noting that the principles used in the proof for Theorem 6.3.15 do not apply in this case. Although, X is a code, the idea of the rewrite rules being extensible does not apply and therefore condition (B) is not fulfilled. To see this we look at the rewrite rules $R = \{(221, 212), (211, 121)\}$ with the alphabet $X_0 = \{1, 2^i\}$. To illustrate the problem which arises when trying to use this theorem we set $i = 3$. Consider the first rewrite rule $(221, 212)$ and recall how we build a new set of rewrite rules, namely R_0 as described in Section 6.3.3. Then a minimal extension s of r^{+1} to $\langle X_0 \rangle$ would be $s \equiv xr^{+1}y \equiv 2221 \in \langle X_0 \rangle$. However, $xr^{-1}y \equiv 2212 \xrightarrow{*}_R 2122 \notin \langle X_0 \rangle$. Therefore $im(R_0) \not\subseteq \langle X_0 \rangle$ and condition (B) is not fulfilled.

Full definitions will follow in this section but first we give an outline of the proof for Proposition 6.5.1. We define a finite presentation $\langle B \mid Q\rangle$ for a monoid M. Then we prove that the associated finite rewriting system (B, Q) is complete. Next we define mappings between the monoid defined by the presentation $\langle B \mid Q\rangle$ and the submonoid E of P_2 generated by the set $X = \{1, 2^i\}$. We prove that all the relations Q also hold in the submonoid E. We prove that a mapping between elements of M and E is a bijection and complete the proof of isomorphism between the presentations for E and M. Hence there exists a presentation for the submonoid E which is a finite complete rewriting system.

Before we proceed further we make a few definitions.

Definition 6.5.2. Let $B = \{a, b\}$ be a new alphabet where we order the generators $a < b$. Define a set of relations Q in $B^* \times B^*$, with i fixed and equal to that in set $X = \{1, 2^i\}$, consisting of all the following pairs:

(0) (bba, bab),

(1) (ba^{i+1}, aba^i),

(2) $\{(ba^{j-1}ba, ba^jb) \ : \ 2 \leq j \leq i\}$.

Let M be the monoid defined by the presentation $\langle B \mid Q\rangle$ and associated rewriting system (B, Q).

Throughout the remainder of subsection 6.5, unless stated otherwise, we will use the following as defined in Proposition 6.5.1 and Definition 6.5.2:

(i) the sets A, B, X, R, Q;

(ii) the integer i;

(iii) the monoids P_2, M and submonoid E.

In addition, let η be the smallest congruence on A^* which contains R and ρ be the smallest congruence on B^* which contains Q.

6.5.3 Proofs regarding the monoid presentation $\langle B \mid Q\rangle$

Lemma 6.5.3. *The rewriting system (B, Q) is noetherian.*

Proof. Recall the alphabet $B = \{a, b\}$ where $a < b$ and we define a shortlex ordering (see Definition 2.6.2) on the set of words B^*. Now consider each of the rewrite rules in Q:

(0) (bba, bab),

(1) (ba^{i+1}, aba^i),

(2) $\{(ba^{j-1}ba, ba^jb) \ : \ 2 \leq j \leq i\}$.

Firstly, they are all length preserving. Secondly, for any arbitrary rewrite rule with $u, v \in B^*$ and $(u, v) \in Q$ we have $u > v$. This can be seen as in each rule, a letter a is moved one position to the left, exchanging places with a letter b. So we have a set of rules which is reducing in terms of shortlex ordering, which is an admissible partial ordering by Definition 2.6.2 part (d). Hence by Theorem 2.6.7 the reduction relation $\rightarrow_Q$ on B^* is noetherian. Therefore the rewriting system (B, Q) is noetherian. □

Lemma 6.5.4. *The rewriting system* (B, Q) *is locally confluent.*

Proof. The test for local confluence looks for the resolution of all critical pairs of rewrite rules, see Lemma 2.6.9. If this can be achieved then the rewriting system is locally confluent. Recall that critical pairs of rewrite rules are where the left hand side of two rules overlap.

We will now examine the rewrite rules in Q, looking for and resolving all possible critical pairs. For example, all of the rewrite rules numbered (1) and in the set (2) overlap with rule (0). We will use the rules (j) and (k) to represent rules in the set (2). Note that $j, k \in \mathbb{N}$ are variable and we have used both where two rules from the set (2) can overlap. However, i is fixed for the presentation and is equal to that in the set $X = \{1, 2^i\}$. Critical pairs are resolved as follows:

(a) Let $t_1 \equiv bba^{i+1}$, which arises from overlapping the left hand sides of (0) and (1). Then
apply (0) $t_1 \xrightarrow{(0)} baba^i$,
apply (1) $t_1 \xrightarrow{(1)} baba^i$.

(b) Let $t_2 \equiv bba^{j-1}ba$, which arises from overlapping the left hand sides of (0) and (2). Then
apply (0) $t_2 \xrightarrow{(0)} baba^{j-2}ba \xrightarrow{(j-1)} baba^{j-1}b$,
apply (j) $t_2 \xrightarrow{(j)} bba^{j}b \xrightarrow{(0)} baba^{j-1}b$.

(c) Let $t_3 \equiv ba^{j-1}ba^{i+1}$,which arises from overlapping the left hand sides of (2) and (1). Then
apply (j) $t_3 \xrightarrow{(j)} ba^{j}ba^{i}$,
apply (1) $t_3 \xrightarrow{(1)} ba^{j-1}aba^{i}$.

(d) Let $t_4 \equiv ba^{j-1}ba^{k-1}ba$ with $2 \leq k < j \leq i$, which arises from overlapping the left hand sides of (2) and (2). Then
apply (j) $t_4 \xrightarrow{(j)} ba^{j}ba^{k-2}ba$ then
if $k = 2$ $ba^{j}ba^{k-2}ba \equiv ba^{j}bba \xrightarrow{(0)} ba^{j}ba^{k-1}b$ or
if $k > 2$ $\xrightarrow{(k-1)} ba^{j}ba^{k-1}b$,
apply (k) $t_4 \xrightarrow{(k)} ba^{j-1}ba^{k}b \xrightarrow{(j)} ba^{j}ba^{k-1}b$.

(e) Let $t_5 \equiv ba^{k-1}ba^{j-1}ba$ where $2 \leq k < j \leq i$, which arises from overlapping the left hand sides of (2) and (2). Then
apply (k) $t_5 \xrightarrow{(k)} ba^{k}ba^{j-2}ba \xrightarrow{(j-1)} ba^{k}ba^{j-1}b$,
apply (j) $t_5 \xrightarrow{(j)} ba^{k-1}ba^{j}b \xrightarrow{(k)} ba^{k}ba^{j-1}b$.

(f) Let $t_6 \equiv ba^{j-1}ba^{j-1}ba$ where $j \leq i$, which arises from overlapping the left hand sides of (2) and (2). Then
apply first (j) $t_6 \xrightarrow{(j)} ba^{j}ba^{j-2}ba \xrightarrow{(j-1)} ba^{j}ba^{j-1}b$,
apply second (j) $t_6 \xrightarrow{(j)} ba^{j-1}ba^{j}b \xrightarrow{(j)} ba^{j}ba^{j-1}b$.

All critical pairs resolve, so the rewriting system (B, Q) is locally confluent. □

Lemma 6.5.5. *The presentation $\langle B \mid Q \rangle$ for the monoid M is a finite complete rewriting system.*

Proof. The rewriting system (B, Q) for the monoid M is noetherian and locally confluent by Lemmas 6.5.3 and 6.5.4 respectively. So by Lemma 2.6.10 it is also complete. Thus (B, Q) is a finite complete rewriting system. □

Lemma 6.5.6. *The irreducible words with respect to the rewriting system* (B, Q) *are precisely the words of the form:*

$$(a)^p(ba^i)^r(ba^h)^{s_h}(b)^q$$

with integers $p, r, q \geq 0$, $s_h \in \{0, 1\}$ *and* $0 < h < i$ *where* i *is fixed and equal to that in the set* $X = \{1, 2^i\}$*. Hence this set of words gives a set of normal forms for the elements of the monoid* M *defined by the presentation* $\langle B \mid Q \rangle$*.*

Proof. The rewriting system (B, Q) for the monoid M is noetherian and locally confluent by Lemmas 6.5.3 and 6.5.4 respectively. Therefore by Theorem 2.6.13 each equivalence class under $\overset{*}{\leftrightarrow}_Q$ contains a unique irreducible word, which we will define with the above normal form. To prove that we have the correct normal form we intend to show that it is irreducible and that every word can be reduced to one in this form.

First we prove that the normal form is irreducible. To do this we show that it does not contain any substring of letters which also occur as the left hand side of a rewrite rule in the set Q. Note that none of the left hand sides of the rules start with a letter a so we can ignore the first term $(a)^p$ as it cannot be part of a string within a word that could be rewritten. Also, none end in a b so we can ignore the last term $(b)^q$ for a similar reason.

Consider the second term $(ba^i)^r$ with $r \geq 0$. If $r = 1$ then the string ba^i only appears in the left hand side of rule (1) which is the string ba^{i+1}. So, in order to have a string of letters which could be rewritten, then ba^i would have to be followed by an a. This is not possible as it can only be followed by a b i.e. another of this term if $r > 1$ or the 3rd or 4th term. Also, no part of the a^i can be the prefix to a string that occurs in a rule as none begin with an a. So the second term cannot be part of a string within a word that could be rewritten.

Next we look at the third term $(ba^h)^{s_h}$ which varies depending on the value of h, although if it exists it is defined as $0 < h < i$ with i fixed. As we have already discounted the first and second terms as any potential prefix to a string of letters in a rule, then we need to consider this term as a whole or as a proper prefix. Looking at the left hand side of the rules we can see that the third term could potentially be the prefix to any of the rules in the set numbered (2). However, to achieve a complete matching string it would have to be followed by ba. This is not possible as the only term that can follow is the final term $(b)^q$. So the third term cannot be part of a string within a word that could be rewritten.

In summary, having considered all possible combinations of terms in the proposed normal form, there does not exist a substring of letters which match the left hand side of a rewrite rule. Therefore, the proposed normal form is irreducible.

Next we aim to show that every word in B^* can be reduced to a word in the proposed normal form. In other words, if any word, say $t \in B^*$ does not contain the left hand side string of letters from any rewrite rule (i.e. it is irreducible), then t must be in the proposed normal form. To do this we use proof by induction on the length of the word t.

Induction statement:

Let $t \in B^*$ be an irreducible word which is in the proposed normal form. Let $t_1 \equiv th$ where $h \in B = \{a, b\}$ so that $|t_1| = |t| + 1$. Next apply the rewrite rules Q to t_1 such that $t_1 \xrightarrow{*}_Q t_1'$ and t_1' is irreducible. Then the irreducible word t_1' will be in the proposed normal form.

Firstly, if $|t| = 0$ then $t_1 \equiv a$ or $t_1 \equiv b$ and both these words are in the proposed normal form. Thus the induction statement is true for $|t| = 0$.

Next we look at every possible form for the word t which is in the proposed normal form, append either an a or a b and apply the rewrite rules Q to the resulting word. Then we check to see if our irreducible word is in the proposed normal form. Note that the value of i is fixed as per the presentation $\langle B \mid Q \rangle$.

Appending a b will clearly always result in an irreducible word that is already in the proposed normal form. So we look at each case where we can append an a, as follows:

(i) Let $t \equiv (a)^p$ with $p > 0$,

then $t_1 \equiv (a)^p a = (a)^{p+1}$ which is in the proposed normal form.

(ii) Let $t \equiv (ba^i)^r$ with $r > 0$,

then $t_1 \equiv (ba^i)^r a \xrightarrow{(1)}{}^{*} (a)^p(ba^i)^r$ with $p = 1$ which is in the proposed normal form.

(iii) Let $t \equiv ba^h$ with $0 < h < i$, then $t_1 \equiv ba^h a \equiv ba^{h+1}$

and if $h + 1 < i$ then $t_1 \equiv ba^{h+1}$

otherwise $h + 1 = i$ and so $t_1 \equiv (ba^i)^r$ with $r = 1$,

both of which are in the proposed normal form.

(iv) Let $t \equiv (b)^q$ with $q > 0$

and if $q = 1$ then $t_1 \equiv ba \equiv (ba^h)$ with $h = 1$

otherwise if $q > 1$ then $t_1 \equiv (b)^q a \xrightarrow{(0)}{}^{*} (ba^h)(b)^{q-1}$ with $h = 1$,

both of which are in the proposed normal form.

(v) Let $t \equiv (a)^p(b)^q$ with $p, q > 0$ and as no rules begin with the letter a we can look at part (iv) which gives two outcomes,

if $q = 1$ then $t_1 \equiv (a)^p ba \equiv (a)^p(ba^h)$ with $h = 1$

otherwise if $q > 1$ then $t_1 \equiv (a)^p(b)^q a \xrightarrow{(0)}{}^{*} (a)^p(ba^h)(b)^{q-1}$ with $h = 1$,

both of which are in the proposed normal form.

(vi) Let $t \equiv (a)^p(ba^i)^r$ with $p, r > 0$ and as no rewrite rule begins with the letter a we can look at part (ii),

then $t_1 \equiv (a)^p(ba^i)^r a \xrightarrow{(1)}{}^{*} (a)^{p+1}(ba^i)^r$ which is in the proposed normal form.

(vii) Let $t \equiv (a)^p(ba^h)$ with $p > 0$ and $0 < h < i$ and see part (iii),

and if $h + 1 < i$ then $t_1 \equiv (a)^p(ba^{h+1})$, otherwise $h+1=i$ and so $t_1 \equiv (a)^p(ba^i)^r$ with $r=1$,

both of which are in the proposed normal form.

(viii) Let $t \equiv (ba^i)^r(b)^q$ with $r,q>0$ and see part (iv),

and if $q=1$ then $t_1 \equiv (ba^i)^r ba \equiv (ba^i)^r(ba^h)$ with $h=1$,

otherwise if $q>1$ then $t_1 \equiv (ba^i)^r(b)^q a \xrightarrow{(0)}{}^{*} (ba^i)^r(ba^h)(b)^{q-1}$

with $h=1$,

both of which are in the proposed normal form.

(ix) Let $t \equiv (ba^i)^r(ba^h)$ with $r>0$ and $0<h<i$ and see part (iii),

then $t_1 \equiv (ba^i)^r(ba^h)a \equiv (ba^i)^r ba^{h+1}$

and if $h+1<i$ then $t_1 \equiv (ba^i)^r(ba^{h+1})$

otherwise $h+1=i$ and so $t_1 \equiv (ba^i)^{r+1}$,

both of which are in the proposed normal form.

(x) Let $t \equiv (ba^h)(b)^q$ with $q>0$ and $0<h<i$ and see parts (iv) and (iii),

if $q=1$ then

$t_1 \equiv (ba^h)ba$ and now the result depends on h

so if $q=1$ and $h+1<i$ then $t_1 \xrightarrow{(h+1)} (ba^{h+1})(b)^q$

and if $q=1$ and $h+1=i$ then $t_1 \xrightarrow{(h+1)} (ba^i)^r(b)^q$ with $r=1$,

but if $q>1$ then $t_1 \xrightarrow{(0)}{}^{*} (ba^h)ba(b)^{q-1}$ and now the result depends on h

so if $q>1$ and $h+1<i$ then $t_1 \xrightarrow{(h+1)} (ba^{h+1})(b)^q$

and if $q>1$ and $h+1=i$ then $t_1 \xrightarrow{(h+1)} (ba^i)^r(b)^q$ with $r=1$,

all of which are in the proposed normal form.

(xi) Let $t \equiv (a)^p(ba^i)^r(ba^h)$ with $p,r>0$ and $0<h<i$ and as no rules begin with the letter a we can see at part (ix) which gives two outcomes,

then $t_1 \equiv (a)^p(ba^i)^r(ba^h)a \equiv (a)^p(ba^i)^r ba^{h+1}$

and if $h+1<i$ then $t_1 \equiv (a)^p(ba^i)^r(ba^{h+1})$,

otherwise $h+1=i$ and so $t_1 \equiv (a)^p(ba^i)^{r+1}$,

both of which are in the proposed normal form.

(xii) Let $t \equiv (a)^p(ba^i)^r(b)^q$ with $p,r,q>0$ and as no rules begin with the letter a we can see part (viii) which gives two outcomes, and if $q = 1$ then $t_1 \equiv (a)^p(ba^i)^r ba \equiv (a)^p(ba^i)^r(ba^h)$ with $h = 1$

otherwise if $q > 1$ then $t_1 \equiv (a)^p(ba^i)^r(b)^q a \xrightarrow{(0)}{}^{*} (a)^p(ba^i)^r(ba^h)(b)^{q-1}$

with $h = 1$,

both of which are in the proposed normal form.

(xiii) Let $t \equiv (a)^p(ba^h)(b)^q$ with $p, q > 0$ and $0 < h < i$ and we can see part (x) which gives various outcomes,

if $q = 1$ then

$t_1 \equiv (a)^p(ba^h)ba$ and now the result depends on h

so if $q = 1$ and $h + 1 < i$ then $t_1 \xrightarrow{(h+1)} (a)^p(ba^{h+1})(b)^q$

and if $q = 1$ and $h + 1 = i$ then $t_1 \xrightarrow{(h+1)} (a)^p(ba^i)^r(b)^q$ with $r = 1$,

but if $q > 1$ then $t_1 \xrightarrow{(0)}{}^{*} (a)^p(ba^h)ba(b)^{q-1}$ and the result depends on h

so if $q > 1$ and $h + 1 < i$ then $t_1 \xrightarrow{(h+1)} (a)^p(ba^{h+1})(b)^q$

and if $q > 1$ and $h + 1 = i$ then $t_1 \xrightarrow{(h+1)} (a)^p(ba^i)^r(b)^q$ with $r = 1$,

all of which are in the proposed normal form.

(xiv) Let $t \equiv (ba^i)^r(ba^h)(b)^q$ with $p, r, q > 0$ and $0 < h < i$ and we can see parts (xiii) and (ix) for various outcomes depending on q and h,

if $q = 1$ then

$t_1 \equiv (ba^i)^r(ba^h)ba$ and now the result depends on h

so if $q = 1$ and $h + 1 < i$ then $t_1 \xrightarrow{(h+1)} (ba^i)^r(ba^{h+1})(b)^q$

and if $q = 1$ and $h + 1 = i$ then $t_1 \xrightarrow{(h+1)} (ba^i)^{r+1}(b)^q$,

but if $q > 1$ then

$t_1 \equiv (ba^i)^r(ba^h)(b)^q a \xrightarrow{(0)}{}^{*} (ba^i)^r(ba^h)ba(b)^{q-1}$ and now the result depends on h

so if $q > 1$ and $h + 1 < i$ then $t_1 \xrightarrow{(h+1)} (ba^i)^r(ba^{h+1})(b)^q$

and if $q > 1$ and $h + 1 = i$ then $t_1 \xrightarrow{(h+1)} (ba^i)^{r+1}(b)^q$,

all of which are in the proposed normal form.

(xv) Let $t \equiv (a)^p(ba^i)^r(ba^h)(b)^q$ with $p, r, q > 0$ and $0 < h < i$ and as no rewrite rules have a left hand side prefixed by an a we can see part (xiv) for various outcomes depending on q and h,

if $q = 1$ and $h + 1 < i$ then $t_1 \xrightarrow{(h+1)} (a)^p(ba^i)^r(ba^{h+1})(b)^q$, if $q = 1$ and $h + 1 = i$ then $t_1 \xrightarrow{(h+1)} (a)^p(ba^i)^{r+1}(b)^q$,

if $q > 1$ and $h + 1 < i$ then

$t_1 \xrightarrow{(0)}^* (a)^p(ba^i)^r(ba^h)ba(b)^{q-1} \xrightarrow{(h+1)} (a)^p(ba^i)^r(ba^{h+1})(b)^q$,

if $q > 1$ and $h + 1 = i$ then

$t_1 \xrightarrow{(0)}^* (a)^p(ba^i)^r(ba^h)ba(b)^{q-1} \xrightarrow{(h+1)} (a)^p(ba^i)^{r+1}(b)^q$,

all of which are in the proposed normal form.

In summary, for every irreducible word in the proposed normal form, if we append a single letter from the alphabet B, we can reduce the word to an irreducible word and that word is also in the proposed normal form. Thus the induction statement is true for words of length $|t| + 1$. It has also been proved true for words where $|t| = 0$. Therefore all words in B^* can be reduced to the proposed normal form.

This completes the proof that the irreducible words with respect to the rewriting system (B, Q) are

$$(a)^p(ba^i)^r(ba^h)^{s_h}(b)^q$$

with integers $p, r, q \geq 0$, $s_h \in \{0, 1\}$ and $0 < h < i$. Hence, this set of words is a normal form for the presentation $\langle B \mid Q \rangle$ for the monoid M. □

6.5.4 Proofs regarding the submonoid E

In the previous section we have defined a monoid M which has presentation $\langle B \mid Q \rangle$ which is a finite complete rewriting system. Our aim now is to prove that the submonoid E is isomorphic to the monoid defined by the presentation $\langle B \mid Q \rangle$ and hence E admits a presentation by a finite complete rewriting system, completing the proof of Theorem 6.5.14. To do this we create mappings between elements in M and E and prove that we have a bijection. We also prove that all the relations in Q also hold when mapped to the submonoid E. First some definitions and a theorem which we will make use of in the proof.

The following theorem is courtesy of the notes for a UEA course on Semigroups.

It brings together generally known facts relating to semigroups which we can apply to our proof of the isomorphism between the monoids M and E. Further information can also be found in [47, Chapter 1, Section 2].

Theorem 6.5.7 (Theorem 6.5)**.** *(I) The monoid $S = A^*/\rho$ defined by the presentation $\langle A \mid R \rangle$ is generated by the classes a/ρ ($a \in A$), and these generators satisfy all the defining relations from R.*

(II) Let T be any monoid and assume that there exists a mapping $f : A \to T$ onto a generating set of T, and let $\phi : A^* \to T$ be the unique extension of f to a homomorphism. If the generators Af of T satisfy all the relations from R, then the mapping $\psi : S \to T$ defined by $(w/\rho)\psi = w\phi$ is a well defined epimorphism. In particular, T is a homomorphic image of S.

Definition 6.5.8. Define a mapping $\phi_1 : B \to E$ such that $a \mapsto 1/\eta$ and $b \mapsto 2^i/\eta$. This extends to the unique surjective homomorphism $\phi : B^* \to E$ such that if $t, t_1 \in B^*$, $t_2 \in B$ and $t \equiv t_1 t_2$ then $t\phi = (t_1\phi)t_2\phi$.

Next we prove that all the relations in the monoid M also hold when mapped to the submonoid E. This will enable us to make use of Theorem 6.5.7 part (II).

Lemma 6.5.9. *If $(u, v) \in Q$ is an arbitrary relation then $u\phi = v\phi$ in the submonoid E.*

Proof. Let $u\phi = u'/\eta$ and $v\phi = v'/\eta$. In order to prove that $u'/\eta = v'/\eta$ we will look to prove that the tableau words are the same, that is $P(u') \equiv P(v')$. As the tableau word for an element is unique we will have proved that $u\phi = v\phi$. Note that an unconventional format has been adopted for the Young tableau diagrams where we write 2^i in one box to denote a row of i boxes each with an entry 2. This is in order to handle the multiple occurrences of 2's which varies depending on the value of i. We take each relation in turn and apply Schensted's algorithm.

(0) $\big((bba)\phi,\ (bab)\phi \big) = \big((2^i 2^i 1)/\eta,\ (2^i 1 2^i)/\eta \big)$

We have $P(2^i2^i1) \equiv 2 \;\; 12^{i-1}2^i$ since applying Schensted's algorithm to the left hand side of the relation gives:

$$\begin{array}{|c|}\hline 2^i \\ \hline\end{array} \quad \begin{array}{|c|c|}\hline 2^i & 2^i \\ \hline\end{array} \quad \begin{array}{|c|c|c|}\cline{1-1} 2 & \multicolumn{2}{c}{} \\ \hline 1 & 2^{i-1} & 2^i \\ \hline\end{array}$$

Alternatively, we get the same result $P(2^i12^i) \equiv 2 \;\; 12^{i-1}2^i$, if we apply Schensted's algorithm to the right hand side of the relation:

$$\begin{array}{|c|}\hline 2^i \\ \hline\end{array} \quad \begin{array}{|c|c|}\cline{1-1} 2 & \multicolumn{1}{c}{} \\ \hline 1 & 2^{i-1} \\ \hline\end{array} \quad \begin{array}{|c|c|c|}\cline{1-1} 2 & \multicolumn{2}{c}{} \\ \hline 1 & 2^{i-1} & 2^i \\ \hline\end{array}$$

(1) $\big((ba^{i+1})\phi,\ (aba^i)\phi \big) = \big((2^i1^{i+1})/\eta,\ (12^i1^i)/\eta \big)$

We have $P(2^i1^{i+1}) \equiv 2^i \;\; 1^i1$ since applying Schensted's algorithm to the left hand side of the relation gives:

$$\begin{array}{|c|}\hline 2^i \\ \hline\end{array} \quad \begin{array}{|c|c|}\cline{1-1} 2^i & \multicolumn{1}{c}{} \\ \hline 1^i & 1 \\ \hline\end{array}$$

Alternatively, we get the same result $P(12^i1^i) \equiv 2^i \;\; 1^i1$, if we apply Schensted's algorithm to the right hand side of the relation:

$$\begin{array}{|c|}\hline 1 \\ \hline\end{array} \quad \begin{array}{|c|c|}\hline 1 & 2^i \\ \hline\end{array} \quad \begin{array}{|c|c|}\cline{1-1} 2^i & \multicolumn{1}{c}{} \\ \hline 1^i & 1 \\ \hline\end{array}$$

(2) $\big((ba^{j-1}ba)\phi,\ (ba^jb)\phi \big) = \big((2^i1^{j-1}2^i1)/\eta,\ (2^i1^j2^i)/\eta \big)$ with $i \geq j$

We have $P(2^i1^{j-1}2^i1) \equiv 2^j \;\; 1^j2^{i-j}2^i$ since applying Schensted's algorithm to the left hand side of the relation gives:

$$\begin{array}{|c|}\hline 2^i \\ \hline\end{array} \quad \begin{array}{|c|c|}\cline{1-1} 2^{j-1} & \multicolumn{1}{c}{} \\ \hline 1^{j-1} & 2^{i-j+1} \\ \hline\end{array}$$

$$\begin{array}{|c|c|c|}\cline{1-1} 2^{j-1} & \multicolumn{2}{c}{} \\ \hline 1^{j-1} & 2^{i-j+1} & 2^{i} \\ \hline \end{array} \qquad \begin{array}{|c|c|c|}\cline{1-1} 2^{j} & \multicolumn{2}{c}{} \\ \hline 1^{j} & 2^{i-j} & 2^{i} \\ \hline \end{array}$$

Alternatively, we get the same result $P(2^i 1^j 2^i) \equiv 2^j \;\; 1^j 2^{i-j} 2^i$, if we apply Schensted's algorithm to the right hand side of the relation:

$$\begin{array}{|c|}\hline 2^{i} \\ \hline \end{array}\;, \quad \begin{array}{|c|c|}\cline{1-1} 2^{j} & \multicolumn{1}{c}{} \\ \hline 1^{j} & 2^{i-j} \\ \hline \end{array}\;, \quad \begin{array}{|c|c|c|}\cline{1-1} 2^{j} & \multicolumn{2}{c}{} \\ \hline 1^{j} & 2^{i-j} & 2^{i} \\ \hline \end{array}$$

From the tableau above we can see that in each case $u\phi = v\phi$ in P_2 and therefore also in the submonoid E. Hence, all the relations Q hold in E. $\square$

Lemma 6.5.10. *The mapping $\psi : M \to E$ defined by $\psi : M \to E$ such that $(t/\rho)\psi = t\phi$ where $t \in B^*$ and so $t/\rho \in M$, is a well-defined epimorphism.*

Proof. This follows from Definition 6.5.8, Theorem 6.5.7 and Lemma 6.5.9. $\square$

Next we aim to prove that the mapping ψ is injective. To do this we will take two words t_1 and t_2 in B^* that are not equal in M. Taking their normal form, map using ψ to the submonoid E and assume for a contradiction that they map to the same element in E. So, if we first take $(t_1/\rho)\psi = t_1\phi = w_1/\eta$. Then we look to find the normal form in P_2 for w_1/η by reducing w_1 using the rewrite rules R such that $w_1 \xrightarrow{*}_R w_1' \in \mathrm{IRR}(R)$. We determine $w_2' \in \mathrm{IRR}(R)$ from $t_2 \in B^*$ in the same manner. Finally, we can compare w_1' with w_2' for a contradiction to our original assumption. At this point an example may help the reader.

Example 6.5.11. Let $i = 3$ be fixed just for the purposes of this example and $t \in \mathrm{IRR}(Q)$ be a specific word in the monoid M such that $t \equiv (a)(ba^3)^2(ba^2)(b)$. Then write $(t/\rho)\psi = t\phi = w/\eta$ which means that $w \equiv (1)(2^3 1^3)^2(2^3 1^2)(2^3)$. Next we want to reduce w using the rewrite rules in R so that we have a word in normal form for the monoid P_2 as per Corollary 6.4.1. Computing the irreducible word we obtain:

$$\begin{aligned}
w \equiv (1)(2^3 1^3)(2^3 1^3)(2^3 1^2)(2^3) &\equiv (1)222111(222111)(22211)(222) \\
&\rightarrow (1)221211(222111)(22211)(222) \\
&\rightarrow (1)212211(222111)(22211)(222) \\
&\rightarrow (1)(212121)222111(22211)(222) \\
&\xrightarrow{*} (1)(212121212121)22211(222) \\
&\xrightarrow{*} (1)(2121212121212121)(2222) \\
&\equiv 1\ (21)^8\ 2^4\ \equiv w'\ \in \mathrm{IRR}(R).
\end{aligned}$$

Note that $w \in X^*$, w/η and $w'/\eta \in E$, $w/\eta =_{P_2} w'/\eta$ but $w' \notin X^*$. In order to check whether words are mapped using $\psi : M \to E$ to the same element in E we will check whether they map to the same element in P_2, as we will see later.

Lemma 6.5.12. *Let $t \in B^*$ such that $t \equiv (a)^p (ba^i)^r (ba^h)^{s_h} (b)^q \in \mathrm{IRR}(Q)$ with integers $p, r, q \geq 0$, $s_h \in \{0, 1\}$ and $0 < h < i$ which is in normal form for the presentation $\langle B \mid Q \rangle$. Let $(t/\rho)\psi = t\phi = w/\eta$ and $w \xrightarrow{*}_R w' \in \mathrm{IRR}(R)$. Then*

$$w' \equiv (1)^c (21)^d (2)^e \text{ where } c = p,\ d = ir + s_h(h) \text{ and } e = iq + s_h(i - h).$$

Proof. By Lemma 6.5.6 t is in normal form for the presentation $\langle B \mid Q \rangle$. By definition of ϕ we can see that $w \equiv (1)^p (2^i 1^i)^r (2^i 1^h)^{s_h} (2^i)^q$. There are no rewrite rules in R where the left hand side starts with the letter 1 or ends with the letter 2. So we need only consider the two middle terms. We will first look at the term $(2^i 1^i)^r$, recall $i > 1$, and we compute the irreducible word that this term would equal after applying the rewrite rules R. This gives us:

$$2^i1^i \equiv 2^{i-2}(221)1^{i-1} \rightarrow_R 2^{i-2}(212)1^{i-1} \equiv 2^{i-2}(2121)1^{i-2} \tag{6.5.1}$$

$$\equiv 2^{i-3}(221)(21)1^{i-2} \rightarrow_R 2^{i-3}(21)(221)1^{i-2} \tag{6.5.2}$$

$$\rightarrow_R 2^{i-3}(21)(212)1^{i-2} \equiv 2^{i-3}(212121)1^{i-3} \tag{6.5.3}$$

$$\xrightarrow{*}_R (21)^i \in \mathrm{IRR}(R). \tag{6.5.4}$$

Looking at the above sequence we can see that if $i = 2$ then the rewriting sequence stops at line (6.5.1) with $2^21^2 \xrightarrow{*}_R (21)^2 \in \mathrm{IRR}(R)$. Similarly, if $i = 3$ the rewriting sequence stops at line (6.5.3) with $2^31^3 \xrightarrow{*}_R (21)^3 \in \mathrm{IRR}(R)$. Thus we can see that for any $i > 1$ the rewriting sequence stops with $2^i1^i \xrightarrow{*}_R (21)^i \in \mathrm{IRR}(R)$. Hence, if the second term is $(2^i1^i)^r$ the rewriting sequence gives us $(2^i1^i)^r \xrightarrow{*}_R (21)^{ir} \in \mathrm{IRR}(R)$.

Similarly we can look at the next term $(2^i1^h)^{s_h}$ with $0 < h < i$ and assuming $s_h = 1$, recall $s_h \in \{0, 1\}$ by definition (see Lemma 6.5.6). We compute the irreducible word that this term would equal after applying the rewrite rules R.

For $h > 1$ this gives us:

$$2^i1^h \equiv 2^{i-2}(221)1^{h-1} \rightarrow_R 2^{i-2}(212)1^{h-1}$$
$$\equiv 2^{i-2}(2121)1^{h-2} \xrightarrow{*}_R 2^{i-h}(21)^h1^{h-h} \equiv 2^{i-h}(21)^h$$

End of stage (i) which moves 1s left

$$\equiv 2^{i-h-1}(221)(21)^{h-1} \rightarrow_R 2^{i-h-1}(21)(2)(21)^{h-1}$$
$$\equiv 2^{i-h-1}(21)(221)(21)^{h-2} \rightarrow_R 2^{i-h-1}(21)(212)(21)^{h-2}$$
$$\equiv 2^{i-h-1}(21)^2(2)(21)^{h-2}$$
$$\xrightarrow{*}_R 2^{i-h-1}(21)^h(2)(21)^{h-h} \equiv 2^{i-h-1}(21)^h(2)$$

End of stage (ii) which moves a spare 2 right

Then repeat stage (ii), moving spare 2s right

$$\xrightarrow{*}_R 2^{i-h-m}(21)^h(2)^m \quad \text{once } m = i-h \text{ we get}$$
$$\xrightarrow{*}_R (21)^h(2)^{i-h} \in \mathrm{IRR}(R).$$

For $h = 1$ the rewriting sequence is $2^i1 \equiv 2^{i-2}221 \rightarrow 2^{i-2}212 \equiv 2^{i-h-1}(21)^h2 \xrightarrow{*}_R 2^{i-1-m}(21)^h(2)^m \xrightarrow{*}_R (21)^h(2)^{i-h} \equiv (21)2^{i-1} \in \mathrm{IRR}(R)$ where $0 < m \leq i-h$.

Next we can put the rewritten terms together and see if we can apply any further rewrite rules. We consider each combination in turn:
If $w \equiv (1)^p(2^i)^q \equiv 1^p2^{iq} \in \mathrm{IRR}(R)$.
If $w \equiv (1)^p(2^i1^i)^r(2^i)^q \xrightarrow{*}_R 1^p(21)^{ir}2^{iq} \in \mathrm{IRR}(R)$.
If $w \equiv (1)^p(2^i1^h)(2^i)^q \xrightarrow{*}_R 1^p(21)^h(2)^{i-h}2^{iq} \equiv 1^p(21)^h(2)^{iq+i-h} \in \mathrm{IRR}(R)$.
If $w \equiv (1)^p(2^i1^i)^r(2^i1^h)(2^i)^q \xrightarrow{*}_R 1^p(21)^{ir}(21)^h(2)^{i-h}2^{iq} \equiv 1^p(21)^{ir+h}(2)^{iq+i-h} \in \mathrm{IRR}(R)$.

Now we can complete the proof and determine w', as follows:

$$w' \equiv (1)^p(21)^{ir}(21)^{s_h(h)}(2)^{s_h(i-h)}(2)^{iq} \equiv 1^p\ (21)^{ir+s_h(h)}\ 2^{iq+s_h(i-h)},$$

where w' is in $\mathrm{IRR}(R)$ and in normal form for P_2, by Corollary 6.4.1. □

Lemma 6.5.13. *The mapping $\psi : M \to E$ is a well-defined bijection.*

Proof. By Lemma 6.5.10 the mapping ψ is a well-defined epimorphism, that is a surjective homomorphism. Therefore it remains to prove that ψ is an injective mapping. The proof will be by contradiction and we will assume that ψ is not injective.

Let $t_1, t_2 \in B^*$ and our assumption (for a contradiction) is that $t_1/\rho \neq t_2/\rho$ in the monoid M but $(t_1/\rho)\psi = (t_2/\rho)\psi$ in the submonoid E. Next, without loss of generality, let t_1 and t_2 be in normal form as any word in B^* can be reduced to a normal form by Lemma 6.5.6. Let

$$t_1 \equiv (a)^p (ba^i)^r (ba^h)^{s_h} (b)^q \text{ and } t_2 \equiv (a)^{p'} (ba^i)^{r'} (ba^{h'})^{s'_h} (b)^{q'}$$

with integers $p, p', r, r', q, q' \geq 0$; s_h and $s'_h \in \{0, 1\}$; $0 < h < i$ and $0 < h' < i$.

Let $(t_1/\rho)\psi = t_1\phi = w_1/\eta$ and $w_1 \xrightarrow{*}_R w'_1 \in \text{IRR}(R)$. Similarly we will let $(t_2/\rho)\psi = t_2\phi = w_2/\eta$ and $w_2 \xrightarrow{*}_R w'_2 \in \text{IRR}(R)$. By our assumption, let $w'_1/\eta = w'_2/\eta$. As w'_1 and w'_2 are in normal form, this implies that $w'_1 \equiv w'_2$.

By Lemma 6.5.12 we have

$$w'_1 \equiv 1^p\ (21)^{ir+s_h(h)}\ 2^{iq+s_h(i-h)} \text{ and } w'_2 \equiv 1^{p'}\ (21)^{ir'+s'_h(h')}\ 2^{iq'+s'_h(i-h')}.$$

As we have $w'_1 \equiv w'_2$ it follows that we must have $p = p'$. When we consider that $h < i$ and $h' < i$ we must also have $q = q'$, $r = r'$ and therefore $h = h'$ and $s_h = s'_h$. This means that $t_1 \equiv t_2$ and $t_1/\rho = t_2/\rho$ which contradicts our assumption and so ψ is injective, which completes our proof. □

6.5.5 Proof of new theorem

We begin this section by returning to our original proposition and we now have everything in place to prove it as a new theorem.

Theorem 6.5.14. *Let $A = \{1, 2\}$ and $R = \{(221, 212), (211, 121)\}$. Then the plactic monoid P_2 is defined by the presentation $\langle A \mid R\rangle$.*

Fix $i \in \mathbb{N}$ with $i > 1$ and let E be the submonoid of P_2 generated by $X = \{1, 2^i\}$. Set $B = \{a, b\}$ and let Q be the subset of $B^* \times B^*$ consisting of all the following pairs:
(i) (bba, bab),
(ii) (ba^{i+1}, aba^i),
(iii) $\{(ba^{j-1}ba, ba^jb) \ : \ 2 \leq j \leq i\}$.
Then (B, Q) is a finite complete rewriting system defining the monoid E where a and b correspond to the generators 1 and 2^i, respectively.

Proof. In Lemma 6.5.5 we proved that the presentation $\langle B \mid Q\rangle$ defining the monoid M is a FCRS. In Definition 6.5.8 and Lemma 6.5.10 we have defined mappings $\phi : B^* \to E$ and $\psi : M \to E$.

The presentation $\langle B \mid Q\rangle$ is generated by the classes b/ρ where $b \in B$ and ψ is a one to one mapping to the classes $(b/\rho)\psi = x/\eta$ where $x \in X$, which generate the submonoid E. By Lemma 6.5.9 all the relations in Q also hold in E. Also, in Lemma 6.5.13 we prove that ψ is a well-founded bijective homomorphism, in other words an isomorphism between M and E.

As the presentation $\langle B \mid Q\rangle$ which defines the monoid M is a FCRS and we have an isomorphism between M and E, then the presentation $\langle B \mid Q\rangle$ is a finite complete rewriting system which defines the monoid E. $\square$

Corollary 6.5.15. *Let E_{NF} be the set of words defined as follows:*

$$(a)^p(ba^i)^r(ba^h)^{s_h}(b)^q$$

with integers $p, r, q \geq 0$, $s_h \in \{0, 1\}$ and $0 < h < i$ with a fixed i which is equal to that in the set $X = \{1, 2^i\}$.

Then $E_{NF} = \mathrm{IRR}(Q)$ and is a set of normal forms for the submonoid E defined *by the presentation $\langle B \mid Q\rangle$, where a and b in B correspond to the generators* 1 *and* 2^i *in* X*, respectively.*

Proof. By Theorem 6.5.14 $\langle B \mid Q\rangle$ is a presentation which defines the monoid E. The presentation $\langle B \mid Q\rangle$ has a normal form as per Lemma 6.5.6 which is the set $\mathrm{IRR}(Q)$. Therefore $E_{NF} = \mathrm{IRR}(Q)$ and the rest follows. □

Corollary 6.5.16. *Let* $A = \{1, 2, 3\}$ *and*

$$R = \{(332, 323), (322, 232), (331, 313), (311, 131), (221, 212), (211, 121), (231, 213),$$

$$(312, 132), (3212, 2321), (32131, 31321), (32321, 32132)\}.$$

Then the plactic monoid P_3 *is defined by the homogeneous presentation* $\langle A \mid R\rangle$*. Let* E *be the submonoid of* P_3 *generated by* $X = \{k, l^i\}$ *for some fixed* $i \in \mathbb{N}$ *with* $i > 1$ *and* $k, l \in A$ *with* $k < l$*. Set* $B = \{a, b\}$ *and let* Q *be the subset of* $B^* \times B^*$ *consisting of all the following pairs:*

(i) (bba, bab),

(ii) (ba^{i+1}, aba^i),

(iii) $\{(ba^{j-1}ba, ba^jb) \ : \ 2 \leq j \leq i\}$.

Then (B, Q) *is a finite complete rewriting system defining the monoid* E *where* a *and* b *correspond to the generators* k *and* l^i*, respectively.*

Proof. By Theorem 6.5.14, if $k = 1$ and $l = 2$, then $\langle B \mid Q\rangle$ is a FCRS defining the submonoid E. It can be shown that the proof of Theorem 6.5.14 holds for $X = \{1, 3^i\}$ and $X = \{2, 3^i\}$. □

6.6 Submonoid of the plactic monoid P_2 generated by the set $X' = \{1^i, 2\}$

Corollary 6.6.1. *Let* $A = \{1, 2\}$ *and* $R = \{(221, 212), (211, 121)\}$*. Then the plactic monoid* P_2 *is defined by the presentation* $\langle A \mid R\rangle$*.*

Fix $i \in \mathbb{N}$ with $i > 1$ and let E' be the submonoid of P_2 generated by $X' = \{1^i, 2\}$. Set $B = \{a, b\}$ and let H be the subset of $B^* \times B^*$ consisting of all the following pairs:

(i) (baa, aba),

(ii) $(b^{i+1}a, b^i ab)$,

(iii) $\{(bab^{j-1}a, ab^j a) : 2 \leq j \leq i\}$.

Then (B, H) is a finite complete rewriting system defining the monoid E' where a and b correspond to the generators 1^i and 2, respectively.

Proof. The proof is the dual of that for Theorem 6.5.14. □

Corollary 6.6.2. *Let E'_{NF} be the set of words defined as follows:*

$$(a)^p (b^h a)^{s_h} (b^i a)^r (b)^q$$

with integers $p, r, q \geq 0$, $s_h \in \{0, 1\}$ and $0 < h < i$.

Then $E'_{NF} = \mathrm{IRR}(H)$ and is a set of normal forms for the submonoid E' defined by the presentation $\langle B \mid H \rangle$, where a and b in B correspond to the generators 1^i and 2 in X', respectively.

Proof. The proof is the dual of that for Corollary 6.5.15. □

Corollary 6.6.3. *Let $A = \{1, 2, 3\}$ and*

$$R = \{(332, 323), (322, 232), (331, 313), (311, 131), (221, 212), (211, 121), (231, 213),$$

$$(312, 132), (3212, 2321), (32131, 31321), (32321, 32132)\}.$$

Then the plactic monoid P_3 is defined by the homogeneous presentation $\langle A \mid R \rangle$. Let E' be the submonoid of P_3 generated by $X' = \{k^i, l\}$ for some fixed $i \in \mathbb{N}$ with $i > 1$ and $k, l \in A$ with $k < l$. Set $B = \{a, b\}$ and let H be the subset of $B^* \times B^*$ consisting of all the following pairs:

(i) (baa, aba),

(ii) $(b^{i+1}a, b^i ab)$,

(iii) $\{(bab^{j-1}a, ab^j a) : 2 \leq j \leq i\}$.

Then (B, H) *is a finite complete rewriting system defining the monoid* E' *where* a *and* b *correspond to the generators* k^i *and* l*, respectively.*

Proof. By Corollary 6.6.1, if $k = 1$ and $l = 2$, then $\langle B \mid H \rangle$ is a FCRS defining the submonoid E'. It can be shown that the proof of Corollary 6.6.1 holds for $X' = \{1^i, 3\}$ and $X' = \{2^i, 3\}$. □

6.7 Submonoid of the plactic monoid P_2 generated by the set $Y_1 = \{1^2, 2^3\}$

Research in this section continues the theme of considering submonoids of the plactic monoid on two variables i.e. the plactic monoid P_2. These submonoids will be those generated by the finite set $Y = \{1^i, 2^j\}$ where $i, j \in \mathbb{N}$ such that $i \leq j$. We will first look at a result for specific fixed values for i and j before giving a partial conjecture for the submonoid generated by the set $Y = \{1^i, 2^j\}$.

Proposition 6.7.1. *Let* $A = \{1, 2\}$ *and* $R = \{(221, 212), (211, 121)\}$*. Then the plactic monoid* P_2 *is defined by the presentation* $\langle A \mid R \rangle$*.*

Let F_1 *be the submonoid of* P_2 *generated by* $Y_1 = \{1^2, 2^3\}$*. Set* $B = \{a, b\}$ *and let* Q *be the subset of* $B^* \times B^*$ *consisting of all the following pairs:*

(i) $(baaa, abaa)$,

(ii) (bba, bab),

(iii) $(baabaa, ababaa)$,

(iv) $(bababa, babaab)$.

Then (B, Q) *is a finite complete rewriting system defining the monoid* F_1 *where* a *and* b *correspond to the generators* 1^2 *and* 2^3*, respectively.*

The proof follows the same approach as that for Theorem 6.5.14 in Subsection 6.5. Before we begin this proof we make a few definitions.

Definition 6.7.2. Let $B = \{a, b\}$ be a new alphabet where we order the generators $a < b$. Define a set of relations Q in $B^* \times B^*$ consisting of all the following pairs:

(1) $(baaa, abaa)$,

(2) (bba, bab),

(3) $(baabaa, ababaa)$,

(4) $(bababa, babaab)$.

Let M be the monoid defined by the presentation $\langle B \mid Q \rangle$ and associated rewriting system (B, Q).

Throughout the remainder of subsection 6.7, unless stated otherwise, we will use the following as defined in Proposition 6.7.1 and Definition 6.7.2:

(i) the sets A, B, Y_1, R, Q;

(ii) the monoids P_2, M, and the submonoid F_1.

In addition, let η be the smallest congruence on A^* which contains R and ρ be the smallest congruence on B^* which contains Q.

6.7.1 Proofs regarding the monoid presentation $\langle B \mid Q \rangle$

Lemma 6.7.3. *The rewriting system (B, Q) is noetherian.*

Proof. Recall the alphabet $B = \{a, b\}$ where $a < b$ and we define a shortlex ordering on the set of words B^*. Now consider each of the rewrite rules in Q:

(1) $(baaa, abaa)$,

(2) (bba, bab),

(3) $(baabaa, ababaa)$,

(4) $(bababa, babaab)$.

Firstly, they are all length preserving. Secondly, for any arbitrary rewrite rule with $u, v \in B^*$ and $(u, v) \in Q$ we have $u > v$. This can be seen as in each rule, a letter a is moved one position to the left, exchanging places with a letter b. So we have a set of rules which is reducing in terms of shortlex ordering, which is an admissible partial ordering by Definition 2.6.2 part (d). Hence by Theorem 2.6.7 the reduction relation $\rightarrow_Q$ on B^* is noetherian. Therefore the rewriting system (B, Q) is noetherian. $\square$

Lemma 6.7.4. *The rewriting system* (B, Q) *is locally confluent.*

Proof. The test for local confluence looks for the resolution of all critical pairs of rewrite rules, see Lemma 2.6.9. Recall that critical pairs of rewrite rules are where the left hand side of two rules overlap. If this can be achieved then the rewriting system is locally confluent.

We will now examine the rewrite rules in Q, looking for and resolving all possible critical pairs as follows:

(a) Let $t_1 \equiv bbaaa$, which arises from overlapping the left hand sides of (1) and (2). Then
apply (2) $t_1 \xrightarrow{(2)} babaa$,
apply (1) $t_1 \xrightarrow{(1)} babaa$.

(b) Let $t_2 \equiv baabaaa$, which arises from overlapping the left hand sides of (1) and (3). Then
apply (3) $t_2 \xrightarrow{(3)} ababaaa \xrightarrow{(1)} abaabaa$,
apply (1) $t_2 \xrightarrow{(1)} baaabaa \xrightarrow{(1)} abaabaa$.

(c) Let $t_2 \equiv bababaaa$, which arises from overlapping the left hand sides of (1) and (4). Then
apply (4) $t_2 \xrightarrow{(4)} babaabaa$,
apply (1) $t_2 \xrightarrow{(1)} babaabaa$.

(d) Let $t_2 \equiv bbaabaa$, which arises from overlapping the left hand sides of (2) and (3). Then apply (2) $t_2 \xrightarrow{(2)} bababaa$,

apply (3) $t_2 \xrightarrow{(3)} bababaa$.

(e) Let $t_2 \equiv bbababa$, which arises from overlapping the left hand sides of (2) and (4). Then

apply (2) $t_2 \xrightarrow{(2)} babbaba \xrightarrow{(2)} bababba \xrightarrow{(2)} bababab$,

apply (4) $t_2 \xrightarrow{(4)} bbabaab \xrightarrow{(2)} babbaab \xrightarrow{(2)} bababab$.

(f) Let $t_2 \equiv babababaa$, which arises from overlapping the left hand sides of (4) and (3). Then

apply (4) $t_2 \xrightarrow{(4)} babaababaa$,

apply (3) $t_2 \xrightarrow{(3)} babaababaa$.

(g) Let $t_2 \equiv baabaabaa$, which arises from overlapping the left hand sides of (3) and (3). Then

apply (3) to the left $t_2 \xrightarrow{(3)} ababaabaa \xrightarrow{(3)} abaababaa$,

apply (3) to the right $t_2 \xrightarrow{(3)} baaababaa \xrightarrow{(1)} abaababaa$.

(h) Let $t_2 \equiv bababab a$, which arises from overlapping the left hand sides of (4) and (4). Then

apply (4) to the left $t_2 \xrightarrow{(4)} babaabba \xrightarrow{(2)} babaabab$,

apply (4) to the right $t_2 \xrightarrow{(4)} babab aab \xrightarrow{(4)} babaabab$.

(i) Let $t_2 \equiv bababababa$, which arises from overlapping the left hand sides of (4) and (4). Then

apply (4) to the left $t_2 \xrightarrow{(4)} babaabbaba \xrightarrow{(2)} babaababba \xrightarrow{(2)} babaababab$,

apply (4) to the right $t_2 \xrightarrow{(4)} babababaab \xrightarrow{(4)} babaabbaab \xrightarrow{(2)} babaababab$.

All critical pairs resolve, so the rewriting system (B, Q) is locally confluent. □

Lemma 6.7.5. *The presentation $\langle B \mid Q \rangle$ for the monoid M is a finite complete rewriting system.*

Proof. The rewriting system (B, Q) for the monoid M is noetherian and locally confluent by Lemmas 6.7.3 and 6.7.4 respectively. So by Lemma 2.6.10 it is also complete. Thus (B, Q) is a finite complete rewriting system. □

Lemma 6.7.6. *The irreducible words with respect to the rewriting system* (B, Q) *are precisely the words of the form:*

$$(a)^p(baa)^{s_3}(babaa)^r(ba)^{s_2}(b)^q$$

with integers $p, q, r \geq 0$*,* $s_3 \in \{0, 1\}$ *and* $s_2 \in \{0, 1, 2\}$*. Hence this set of words gives a set of normal forms for the elements of the monoid* M *defined by the presentation* $\langle B \mid Q \rangle$*.*

Proof. The rewriting system (B, Q) for the monoid M is noetherian and locally confluent by Lemmas 6.7.3 and 6.7.4 respectively. Therefore by Theorem 2.6.13 each equivalence class under $\overset{*}{\leftrightarrow}_Q$ contains a unique irreducible word, which we will define with the above normal form. To prove that we have the correct normal form we intend to show that it is irreducible and that every word can be reduced to one in this form.

First we prove that the normal form is irreducible. To do this we show that it does not contain any substring of letters which also occur as the left hand side of a rewrite rule in the set Q. Note that none of the left hand sides of the rules start with a letter a so we can ignore the first term $(a)^p$ as it cannot be part of a string within a word that could be rewritten. Also, none end in a b so we can ignore the last term $(b)^q$ for a similar reason.

If the second term $(baa)^{s_3}$ exists we consider the string of letters if it is followed by either the third or fourth term, the latter occurring once or twice. So we have three possible strings *baababaa*, *baaba* or *baababa* and none contain a string of letters from the left hand side of a rewrite rule. Next we look at possible strings that start with the third term $(babaa)^r$. We could have $r > 1$ or $r = 1$ followed by the fourth term, the latter occurring once or twice. So we consider *babaababaa*, *babaaba* and *babaababa* none of which contain a string of letters from the left hand side of a rewrite rule.

In summary, having considered all possible combinations of terms in the proposed normal form, there does not exist a substring of letters which match the left hand side of a rewrite rule. Therefore, the proposed normal form is irreducible.

Next we aim to show that every word in B^* can be reduced to a word in the proposed normal form. In other words, if any word, say $t \in B^*$ does not contain the left hand side string of letters from any rewrite rule (i.e. it is irreducible), then t must be in the proposed normal form. To do this we use proof by induction on the length of the word t.

<u>Induction statement</u>:

Let $t \in B^*$ be an irreducible word which is in the proposed normal form. Let $t_1 \equiv th$ where $h \in B = \{a, b\}$ so that $|t_1| = |t| + 1$. Next apply the rewrite rules Q to t_1 such that $t_1 \xrightarrow{*}_Q t_1'$ and t_1' is irreducible. Then the irreducible word t_1' will be in the proposed normal form.

Firstly, if $|t| = 0$ then $t_1 \equiv a$ or $t_1 \equiv b$ and both these words are in the proposed normal form. Thus the induction statement is true for $|t| = 0$.

Next we look at every possible form for the word t which is in the proposed normal form, append either an a or a b and apply the rewrite rules Q to the resulting word. Then we check to see if our irreducible word is in the proposed normal form. Appending a b will clearly always result in an irreducible word that is already in the proposed normal form. So we look at each case where we can append an a, as follows:

(i) Let $t \equiv (a)^p$ with $p > 0$,

then $t_1 \equiv (a)^p a = (a)^{p+1}$ which is in the proposed normal form.

(ii) Let $t \equiv (baa)^{s_3}$ with $s_3 = 1$,

then $t_1 \equiv (baa)a \xrightarrow{(1)} (a)^p(baa)^{s_3}$ with $p = 1$ which is in the proposed normal form.

(iii) Let $t \equiv (babaa)^r$ with $r > 0$,

then $t_1 \equiv (babaa)^r a \xrightarrow{(1)} (babaa)^{r-1} baabaa \xrightarrow{(3)} (babaa)^{r-1} a(babaa)$ $\xrightarrow{(1,3)}{}^{*}$ $(a)^p (babaa)^r$ with $p = 1$ which is in the proposed normal form.

(iv) Let $t \equiv (ba)^{s_2}$ with $s_2 > 0$,

if $s_2 = 1$ then $t_1 \equiv baa \equiv (baa)^{s_3}$ with $s_3 = 1$,

if $s_2 = 2$ then $t_1 \equiv babaa \equiv (babaa)^r$ with $r = 1$,

both of which are in the proposed normal form.

(v) Let $t \equiv (b)^q$ with $q > 0$,

if $q = 1$ then $t_1 \equiv ba \equiv (ba)^{s_2}$ with $s_2 = 1$,

if $q = 2$ then $t_1 \equiv bba \xrightarrow{(2)} (ba)^{s_2}(b)^{q-1}$ with $s_2 = 1$,

if $q \geq 3$ then $t_1 \equiv (b)^{q-2} bba \xrightarrow{(2)} (b)^{q-2} bab \xrightarrow{(2)}{}^{*} (ba)^{s_2}(b)^{q-1}$ with $s_2 = 1$,

all of which are in the proposed normal form.

(vi) Let $t \equiv (a)^p (baa)^{s_3}$ with $p > 0$, $s_3 = 1$ and see part(ii),

then $t_1 \equiv (a)^p (baa)^{s_3} a \xrightarrow{(1)} (a)^{p+1} (baa)^{s_3}$ which is in the proposed normal form.

(vii) Let $t \equiv (a)^p (babaa)^r$ with $p, r > 0$ and see part(iii),

then $t_1 \equiv (a)^p (babaa)^r a \xrightarrow{(1,3)}{}^{*} (a)^{p+1} (babaa)^r$ which is in the proposed normal form.

(viii) Let $t \equiv (a)^p (ba)^{s_2}$ with $s_2 > 0$ and $p > 0$ and see part (iv),

if $s_2 = 1$ then $t_1 \equiv (a)^p baa \equiv (a)^p (baa)^{s_3}$ with $s_3 = 1$,

if $s_2 = 2$ then $t_1 \equiv (a)^p babaa \equiv (a)^p (babaa)^r$ with $r = 1$,

both of which are in the proposed normal form.

(ix) Let $t \equiv (a)^p (b)^q$ with $q > 0$ and $p > 0$ and see part (v),

if $q = 1$ then $t_1 \equiv (a)^p ba \equiv (a)^p (ba)^{s_2}$ with $s_2 = 1$,

if $q = 2$ then $t_1 \equiv (a)^p bba \xrightarrow{(2)} (a)^p (ba)^{s_2} (b)^{q-1}$ with $s_2 = 1$,

if $q \geq 3$ then $t_1 \equiv (a)^p (b)^{q-2} bba \xrightarrow{(2)} (a)^p (b)^{q-2} bab \xrightarrow{(2)}{}^{*} (a)^p (ba)^{s_2} (b)^{q-1}$ with $s_2 = 1$, all of which are in the proposed normal form.

(x) Let $t \equiv (baa)^{s_3} (babaa)^r$ with $r > 0$ and $s_3 = 1$ and see parts (iii) and (ii), then $t_1 \equiv (baa)^{s_3}(babaa)^r a \xrightarrow{(1,3)}{}^{*} (baa)^{s_3} a (babaa)^r \xrightarrow{(1)} (a)^p (baa)^{s_3}(babaa)^r$ with $p = 1$ which is in the proposed normal form.

(xi) Let $t \equiv (baa)^{s_3}(ba)^{s_2}$ with $s_2 > 0$ and $s_3 = 1$,
if $s_2 = 1$ then $t_1 \equiv baabaa \xrightarrow{(3)} (a)^p(babaa)^r$ with $p = 1$ and $r = 1$,
if $s_2 = 2$ then $t_1 \equiv baababaa \equiv (baa)^{s_3}(babaa)^r$ with $r = 1$,
both of which are in the proposed normal form.

(xii) Let $t \equiv (baa)^{s_3}(b)^q$ with $q > 0$, $s_3 = 1$ and see part(v),
then $t_1 \equiv (baa)^{s_3}(b)^q a \xrightarrow{(2)}{}^{*} (baa)^{s_3}(ba)^{s_2}(b)^{q-1}$ with $s_2 = 1$ which is in the proposed normal form.

(xiii) Let $t \equiv (babaa)^r(ba)^{s_2}$ with $s_2 > 0$, $r > 0$ and see part (iii),
if $s_2 = 1$ then $t_1 \equiv (babaa)^r baa \xrightarrow{(3)}{}^{*} (baa)^{s_3}(babaa)^r$ with $s_3 = 1$,
if $s_2 = 2$ then $t_1 \equiv (babaa)^r babaa \equiv (babaa)^{r+1}$,
both of which are in the proposed normal form.

(xiv) Let $t \equiv (babaa)^r(b)^q$ with $q > 0$, $r > 0$ and see part(v),
then $t_1 \equiv (babaa)^r(b)^q a \xrightarrow{(2)}{}^{*} (babaa)^r(ba)^{s_2}(b)^{q-1}$ with $s_2 = 1$ which is in the proposed normal form.

(xv) Let $t \equiv (ba)^{s_2}(b)^q$ with $s_2 > 0$, $q > 0$ and we see parts (iv) and (v),
if $s_2 = 1$ then $t_1 \equiv ba(b)^q a \xrightarrow{(2)}{}^{*} (ba)^{s_{22}}(b)^{q-1}$ with $s_{22} = 2$,
if $s_2 = 2$ then $t_1 \equiv baba(b)^q a \xrightarrow{(2)}{}^{*} bababa(b)^{q-1} \xrightarrow{(4)} (babaa)^r(b)^q$ with $r = 1$, both of which are in the proposed normal form.

(xvi) Let $t \equiv (a)^p(baa)^{s_3}(babaa)^r$ with $p, r > 0$, $s_3 = 1$ and see part (x),
then $t_1 \equiv (a)^p(baa)^{s_3}(babaa)^r a \xrightarrow{(1,3)}{}^{*} (a)^p(baa)a(babaa)^r$
$\xrightarrow{(1)} (a)^{p+1}(baa)^{s_3}(babaa)^r$ which is in the proposed normal form.

(xvii) Let $t \equiv (a)^p(baa)^{s_3}(ba)^{s_2}$ with $p, s_2 > 0$, $s_3 = 1$ and see part (xi),
if $s_2 = 1$ then $t_1 \equiv (a)^p baabaa \xrightarrow{(3)} (a)^{p+1}(babaa)^r$ with $r = 1$,
if $s_2 = 2$ then $t_1 \equiv (a)^p(baa)^{s_3}babaa \equiv (a)^p(baa)^{s_3}(babaa)^r$ with $r = 1$,
both of which are in the proposed normal form.

(xviii) Let $t \equiv (a)^p(baa)^{s_3}(b)^q$ with $p, q > 0$, $s_3 = 1$ and see part(xii), then $t_1 \equiv (a)^p(baa)^{s_3}(b)^q a \xrightarrow{(2)}{}^{*} (a)^p(baa)^{s_3}(ba)^{s_2}(b)^{q-1}$ with $s_2 = 1$ which is in the proposed normal form.

(xix) Let $t \equiv (a)^p(babaa)^r(ba)^{s_2}$ with $p, r, s_2 > 0$ and see part (xiii),
if $s_2 = 1$ then $t_1 \equiv (a)^p(babaa)^r baa \xrightarrow{(1,3)}{}^{*} (a)^p(baa)^{s_3}(babaa)^r$ with $s_3 = 1$,
if $s_2 = 2$ then $t_1 \equiv (a)^p(babaa)^r babaa \equiv (a)^p(babaa)^{r+1}$,
both of which are in the proposed normal form.

(xx) Let $t \equiv (a)^p(babaa)^r(b)^q$ with $p, r, q > 0$ and see part(xiv), then $t_1 \equiv (a)^p(babaa)^r(b)^q a \xrightarrow{(2)}{}^{*} (a)^p(babaa)^r(ba)^{s_2}(b)^{q-1}$ with $s_2 = 1$ which is in the proposed normal form.

(xxi) Let $t \equiv (a)^p(ba)^{s_2}(b)^q$ with $s_2 > 0$, $p, q > 0$ and see part (xv),
if $s_2 = 1$ then $t_1 \equiv (a)^p ba(b)^q a \xrightarrow{(2)}{}^{*} (a)^p(ba)^{s_{22}}(b)^{q-1}$ with $s_{22} = 2$,
if $s_2 = 2$ then $t_1 \equiv (a)^p baba(b)^q a \xrightarrow{(2)}{}^{*} (a)^p bababa(b)^{q-1}$
$\xrightarrow{(4)} (a)^p(babaa)^r(b)^q$ with $r = 1$,
both of which are in the proposed normal form.

(xxii) Let $t \equiv (baa)^{s_3}(babaa)^r(ba)^{s_2}$ with $s_2 > 0$, $s_3, r > 0$ and see part (xiii),
if $s_2 = 1$ then $t_1 \equiv (baa)^{s_3}(babaa)^r baa \xrightarrow{(3)}{}^{*} (baa)(baa)(babaa)^r$
$\xrightarrow{(3)} (a)^p(babaa)^{r+1}$ with $p = 1$,
if $s_2 = 2$ then $t_1 \equiv (baa)^{s_3}(babaa)^r babaa \equiv (baa)^{s_3}(babaa)^{r+1}$,
both of which are in the proposed normal form.

(xxiii) Let $t \equiv (baa)^{s_3}(babaa)^r(b)^q$ with $r, q > 0$, $s_3 = 1$ and see part (xiv), then $t_1 \equiv (baa)^{s_3}(babaa)^r(b)^q a \xrightarrow{(2)}{}^{*} (baa)^{s_3}(babaa)^r(ba)^{s_2}(b)^{q-1}$ with $s_2 = 1$ which is in the proposed normal form.

(xxiv) Let $t \equiv (baa)^{s_3}(ba)^{s_2}(b)^q$ with $s_2, q > 0$, $s_3 = 1$ and see part (xv),
if $s_2 = 1$ then $t_1 \equiv (baa)^{s_3} ba(b)^q a \xrightarrow{(2)}{}^{*} (baa)^{s_3}(ba)^{s_{22}}(b)^{q-1}$ with $s_{22} = 2$,
if $s_2 = 2$ then $t_1 \equiv (baa)^{s_3} baba(b)^q a \xrightarrow{(2)}{}^{*} (baa)^{s_3} bababa(b)^{q-1}$
$\xrightarrow{(4)} (baa)^{s_3}(babaa)^r(b)^q$ with $r = 1$,
both of which are in the proposed normal form.

(xxv) Let $t \equiv (babaa)^r(ba)^{s_2}(b)^q$ with $r, q, s_2 > 0$ and see part (xv),

if $s_2 = 1$ then $t_1 \equiv (babaa)^r ba(b)^q a \xrightarrow{(2)}{}^{*} (babaa)^r(ba)^{s_{22}}(b)^{q-1}$ with $s_{22} = 2$,

if $s_2 = 2$ then $t_1 \equiv (babaa)^r baba(b)^q a \xrightarrow{(2)}{}^{*} (babaa)^r bababa(b)^{q-1}$
$\xrightarrow{(4)} (babaa)^{r+1}(b)^q$,

both of which are in the proposed normal form.

(xxvi) Let $t \equiv (a)^p(baa)^{s_3}(babaa)^r(ba)^{s_2}$ with $s_2 > 0$, $p, r > 0$, $s_3 = 1$ and see part (xxii),

if $s_2 = 1$ then $t_1 \equiv (a)^p(baa)^{s_3}(babaa)^r baa \xrightarrow{(3)}{}^{*} (a)^p(baa)(baa)(babaa)^r$
$\xrightarrow{(3)} (a)^{p+1}(babaa)^{r+1}$,

if $s_2 = 2$ then $t_1 \equiv (baa)^{s_3}(babaa)^r babaa \equiv (baa)^{s_3}(babaa)^{r+1}$,

both of which are in the proposed normal form.

(xxvii) Let $t \equiv (a)^p(baa)^{s_3}(babaa)^r(b)^q$ with $p, r, q > 0$, $s_3 = 1$ and see part (xxiv),

then $t_1 \equiv (a)^p(baa)^{s_3}(babaa)^r(b)^q a \xrightarrow{(2)}{}^{*} (a)^p(baa)^{s_3}(babaa)^r(ba)^{s_2}(b)^{q-1}$
with $s_2 = 1$ which is in the proposed normal form.

(xxviii) Let $t \equiv (a)^p(baa)^{s_3}(ba)^{s_2}(b)^q$ with $p, q, s_2 > 0$, $s_3 = 1$ and see part (xxiv),

if $s_2 = 1$ then $t_1 \equiv (a)^p(baa)^{s_3} ba(b)^q a \xrightarrow{(2)}{}^{*} (a)^p(baa)^{s_3}(ba)^{s_{22}}(b)^{q-1}$ with $s_{22} = 2$,

if $s_2 = 2$ then $t_1 \equiv (a)^p(baa)^{s_3} baba(b)^q a \xrightarrow{(2)}{}^{*} (a)^p(baa)^{s_3} bababa(b)^{q-1}$
$\xrightarrow{(4)} (a)^p(baa)^{s_3}(babaa)^r(b)^q$ with $r = 1$,

both of which are in the proposed normal form.

(xxix) Let $t \equiv (a)^p(babaa)^r(ba)^{s_2}(b)^q$ with $r, q, p, s_2 > 0$ and see part (xxv),

if $s_2 = 1$ then

$t_1 \equiv (a)^p(babaa)^r ba(b)^q a \xrightarrow{(2)}{}^{*} (a)^p(babaa)^r(ba)^{s_{22}}(b)^{q-1}$ with $s_{22} = 2$,

if $s_2 = 2$ then $t_1 \equiv (a)^p(babaa)^r baba(b)^q a \xrightarrow{(2)}{}^{*} (a)^p(babaa)^r bababa(b)^{q-1}$
$\xrightarrow{(4)} (a)^p(babaa)^{r+1}(b)^q$,

both of which are in the proposed normal form.

(xxx) Let $t \equiv (baa)^{s_3}(babaa)^r(ba)^{s_2}(b)^q$ with $r, q, s_2 > 0$, $s_3 = 1$ and see part (xxv),

if $s_2 = 1$ then

$t_1 \equiv (baa)^{s_3}(babaa)^r ba(b)^q a \xrightarrow{(2)}{}^{*} (baa)^{s_3}(babaa)^r (ba)^{s_{22}}(b)^{q-1}$ with $s_{22} = 2$,

if $s_2 = 2$ then

$t_1 \equiv (baa)^{s_3}(babaa)^r baba(b)^q a \xrightarrow{(2)}{}^{*} (baa)^{s_3}(babaa)^r bababa(b)^{q-1}$

$\xrightarrow{(4)} (baa)^{s_3}(babaa)^{r+1}(b)^q$,

both of which are in the proposed normal form.

(xxxi) Let $t \equiv (a)^p(baa)^{s_3}(babaa)^r(ba)^{s_2}(b)^q$ with $p, r, q, s_2 > 0$, $s_3 = 1$ and see part (xxx),

if $s_2 = 1$ then $t_1 \equiv (a)^p(baa)^{s_3}(babaa)^r ba(b)^q a$

$\xrightarrow{(2)}{}^{*} (a)^p(baa)^{s_3}(babaa)^r(ba)^{s_{22}}(b)^{q-1}$ with $s_{22} = 2$,

if $s_2 = 2$ then $t_1 \equiv (a)^p(baa)^{s_3}(babaa)^r baba(b)^q a$

$\xrightarrow{(2)}{}^{*} (a)^p(baa)^{s_3}(babaa)^r bababa(b)^{q-1} \xrightarrow{(4)} (a)^p(baa)^{s_3}(babaa)^{r+1}(b)^q$,

both of which are in the proposed normal form.

In summary, for every irreducible word in the proposed normal form, if we append a single letter from the alphabet B, we can reduce the word to an irreducible word and that word is also in the proposed normal form. Thus the induction statement is true for words of length $|t| + 1$. It has also been proved true for words where $|t| = 0$. Therefore all words in B^* can be reduced to the proposed normal form.

This completes the proof that the irreducible words with respect to the rewriting system (B, Q) are

$$(a)^p(baa)^{s_3}(babaa)^r(ba)^{s_2}(b)^q$$

with integers $p, q, r \geq 0$, $s_3 \in \{0, 1\}$ and $s_2 \in \{0, 1, 2\}$. Hence, this set of words is a normal form for the presentation $\langle B \mid Q \rangle$ for the monoid M. □

6.7.2 Proofs regarding the submonoid F_1

In the previous section we have defined a monoid M which has presentation $\langle B \mid Q \rangle$ which is a finite complete rewriting system. Our aim now is to prove that the submonoid F_1 is isomorphic to the monoid defined by the presentation

$\langle B \mid Q\rangle$ and hence F_1 admits a presentation by a finite complete rewriting system, completing the proof of Theorem 6.7.12. To do this we create mappings between elements in M and F_1 and prove that we have a bijection. We also prove that all the relations in Q also hold when mapped to the submonoid F_1. First some definitions.

Definition 6.7.7. Define a mapping $\phi_1 : B \to F_1$ such that $a \mapsto 11/\eta$ and $b \mapsto 222/\eta$. This extends to the unique surjective homomorphism $\phi : B^* \to F_1$ such that if $t, t_1 \in B^*$, $t_2 \in B$ and $t \equiv t_1t_2$ then $t\phi = (t_1\phi)t_2\phi$.

Lemma 6.7.8. *If $(u, v) \in Q$ is an arbitrary relation then $u\phi = v\phi$ in the submonoid F_1.*

Proof. Let $u\phi = u'/\eta$ and $v\phi = v'/\eta$. In order to prove that $u'/\eta = v'/\eta$ we will look to prove that the tableau words are the same, that is $P(u') \equiv P(v')$. As the tableau word for an element is unique we will have proved that $u\phi = v\phi$. We take each relation in turn and apply Schensted's algorithm.

(1) $\big((baaa)\phi, (abaa)\phi\big)$ $\big((2^3 1^2 1^2 1^2)/\eta, (1^2 2^3 1^2 1^2)/\eta\big)$

We have $P(2^3 1^2 1^2 1^2) \equiv 2^3\ \ 1^6$ since applying Schensted's algorithm to the left hand side of the relation gives:

$$\begin{array}{|c|c|c}\hline 2 & 2 & \multicolumn{1}{|c}{} \\ \hline 1 & 1 & \multicolumn{1}{|c|}{2} \\ \hline \end{array} \qquad \begin{array}{|c|c|c|c}\hline 2 & 2 & 2 & \multicolumn{1}{|c}{} \\ \hline 1 & 1 & 1 & \multicolumn{1}{|c|}{1} \\ \hline \end{array} \qquad \begin{array}{|c|c|c|c|c|c|}\hline 2 & 2 & 2 & \multicolumn{3}{|c}{} \\ \hline 1 & 1 & 1 & 1 & 1 & 1 \\ \hline \end{array}$$

Alternatively, we get the same result $P(1^2 2^3 1^2 1^2) \equiv 2^3\ \ 1^6$, if we apply Schensted's algorithm to the right hand side of the relation:

$$\begin{array}{|c|c|c|c|c|}\hline 1 & 1 & 2 & 2 & 2 \\ \hline \end{array} \qquad \begin{array}{|c|c|c|c|c|}\hline 2 & 2 & \multicolumn{3}{|c}{} \\ \hline 1 & 1 & 1 & 1 & 2 \\ \hline \end{array} \qquad \begin{array}{|c|c|c|c|c|c|}\hline 2 & 2 & 2 & \multicolumn{3}{|c}{} \\ \hline 1 & 1 & 1 & 1 & 1 & 1 \\ \hline \end{array}$$

(2) $\big((bba)\phi, (bab)\phi\big) = \big((2^3 2^3 1^2)/\eta, (2^3 1^2 2^3)/\eta\big)$

We have $P(2^3 2^3 1^2) \equiv 2^2\ \ 1^2 2^4$ since applying Schensted's algorithm to the left hand side of the relation gives:

$$\boxed{2}\boxed{2}\boxed{2} \qquad \boxed{2}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\boxed{2} \qquad \begin{array}{l}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\end{array}$$

Alternatively, we get the same result $P(2^31^22^3) \equiv 2^2 \;\; 1^22^4$, if we apply Schensted's algorithm to the right hand side of the relation:

$$\boxed{2}\boxed{2}\boxed{2} \qquad \begin{array}{l}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{2}\end{array} \qquad \begin{array}{l}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\end{array}$$

(3) $\big((baabaa)\phi, (ababaa)\phi\big) = \big((2^31^21^22^31^21^2)/\eta, (1^22^31^22^31^21^2)/\eta\big)$

We have $P(2^31^21^22^31^21^2) \equiv 2^6 \;\; 1^8$ since applying Schensted's algorithm to the left hand side of the relation gives:

$$\begin{array}{l}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{2}\end{array} \qquad \begin{array}{l}\boxed{2}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{1}\boxed{1}\end{array} \qquad \begin{array}{l}\boxed{2}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{2}\boxed{2}\boxed{2}\end{array}$$

$$\begin{array}{l}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{2}\end{array} \qquad \begin{array}{l}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\end{array}$$

Alternatively, we get the same result $P(1^22^31^22^31^21^2) \equiv 2^6 \;\; 1^8$, if we apply Schensted's algorithm to the right hand side of the relation:

$$\boxed{1}\boxed{1}\boxed{2}\boxed{2}\boxed{2} \qquad \begin{array}{l}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{2}\end{array}$$

$$\begin{array}{l}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\end{array} \qquad \begin{array}{l}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{2}\boxed{2}\end{array} \qquad \begin{array}{l}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\end{array}$$

(4) $\big((bababa)\phi, (babaab)\phi\big) = \big((2^31^22^31^22^31^2)/\eta, (2^31^22^31^21^22^3)/\eta\big)$

We have $P(2^31^22^31^22^31^2) \equiv 2^6 \;\; 1^62^3$ since applying Schensted's algorithm to the left hand side of the relation gives:

$$\begin{array}{l}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{2}\end{array} \qquad \begin{array}{l}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\end{array} \qquad \begin{array}{l}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{2}\boxed{2}\end{array}$$

$$\begin{array}{l}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\end{array} \qquad \begin{array}{l}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\boxed{2}\\ \boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{2}\boxed{2}\boxed{2}\end{array}$$

Alternatively, we get the same result $P(2^31^22^31^21^22^3) \equiv 2^6 \;\; 1^62^3$, if we apply Schensted's algorithm to the right hand side of the relation:

$$\begin{array}{ccc} 2&2& \\ 1&1&2 \end{array} \qquad \begin{array}{cccccc} 2&2&&&& \\ 1&1&2&2&2&2 \end{array} \qquad \begin{array}{cccccc} 2&2&2&2&& \\ 1&1&1&1&2&2 \end{array}$$

$$\begin{array}{cccccc} 2&2&2&2&2&2 \\ 1&1&1&1&1&1 \end{array} \qquad \begin{array}{ccccccccc} 2&2&2&2&2&2&&& \\ 1&1&1&1&1&1&2&2&2 \end{array}$$

From the tableau above we can see that in each case $u\phi = v\phi$ in P_2 and therefore also in the submonoid F_1. Hence all the relations Q hold in F_1. □

Lemma 6.7.9. *The mapping $\psi : M \to F_1$ defined by $\psi : M \to F_1$ such that $(t/\rho)\psi = t\phi$ where $t \in B^*$ and so $t/\rho \in M$, is a well-defined epimorphism.*

Proof. This follows from Definition 6.7.7, Theorem 6.5.7 and Lemma 6.7.8. □

Next we aim to prove that the mapping ψ is injective and we follow the same approach as in Subsection 6.5.

Lemma 6.7.10. *Let $t \in B^*$ such that $t \equiv (a)^p(baa)^{s_3}(babaa)^r(ba)^{s_2}(b)^q$ in $\mathrm{IRR}(Q)$ with integers $p, q, r \geq 0$, $s_3 \in \{0,1\}$ and $s_2 \in \{0,1,2\}$ which is a normal form for the presentation $\langle B \mid Q\rangle$. Let $(t/\rho)\psi = t\phi = w/\eta$ and $w \xrightarrow{*}_R w' \in \mathrm{IRR}(R)$. Then*

$$w' \equiv (1)^c(21)^d(2)^e \text{ where } c = 2p + s_3,\ d = 3s_3 + 6r + 2s_2 \text{ and } e = 3q + s_2.$$

Proof. By Lemma 6.7.6 t is in normal form for the presentation $\langle B \mid Q\rangle$. We can see that $w \equiv (11)^p(2221111)^{s_3}(222112221111)^r(22211)^{s_2}(222)^q$ by definition of ϕ. Recall $R = \{(211, 121), (221, 212)\}$. There are no rewrite rules in R where the left hand side starts with the letter 1 or ends with the letter 2. So we need only consider the three middle terms.

We first apply the rewrite rules to the term 2221111 which gives:

$$\begin{aligned}2221111 &\equiv 2(221)111 \to_R (221)2111 \to_R 21(221)11 \to_R 2121(211)\\ &\to_R 21(211)(21) \to_R (211)(2121) \to_R (1)(212121)\\ &\equiv (1)(21)^3 \in \mathrm{IRR}(R).\end{aligned}$$

Then applying the rewrite rules to 222112221111 gives:

$$\begin{aligned}222112221111 &\equiv 2(221)12(221)111 \xrightarrow{*}_R (221)21(221)2111 \xrightarrow{*}_R 21(221)21(221)11\\ &\xrightarrow{*}_R (2121)(221)21211 \to_R (212121)(221)211\\ &\to_R (21212121)(221)1 \to_R (21212121)2121\\ &\equiv (21)^6 \in \mathrm{IRR}(R).\end{aligned}$$

Finally we look at the fourth term $(22211)^{s_2}$ where $s_2 \in \{0,1,2\}$. If $s_2 = 1$ then applying the rewrite rules to 22211 gives:

$$\begin{aligned}(22211) &\equiv 2(221)1 \to_R (221)21 \to_R 21(221)\\ &\to_R (2121)2 \equiv (21)^2 2 \in \mathrm{IRR}(R).\end{aligned}$$

If $s_2 = 2$ then applying the rewire rules to 2221122211 gives:

$$\begin{aligned}(22211)(22211) &\equiv 2(221)1(22211) \to_R (221)21(22211)\\ &\to_R 21(221)(22211) \to_R (2121)2(22211)\\ &\equiv (2121)22(221)1 \to_R (2121)2(221)21 \to_R (2121)(221)221\\ &\to_R (2121)212(221) \to_R (212121)(221)2 \to_R (212121)2122\\ &\equiv (21)^4(2)^2 \in \mathrm{IRR}(R).\end{aligned}$$

Now we can put the rewritten terms together and see if we can apply any further rewrite rules. We consider each combination in turn:

If $w \equiv (11)^p(222)^q \equiv 1^{2p}\ 2^{3q} \in \text{IRR}(R)$.

If $w \equiv (11)^p(2221111)(222)^q \equiv 1^{2p}\ 1\ (21)^3\ 2^{3q} \in \text{IRR}(R)$.

If $w \equiv (11)^p(222112221111)^r(222)^q \equiv 1^{2p}\ (21)^{6r}\ 2^{3q} \in \text{IRR}(R)$.

If $w \equiv (11)^p(22211)^{s_2}(222)^q \equiv 1^{2p}\ (21)^{2s_2}\ 2^{s_2}\ 2^{3q} \in \text{IRR}(R)$.

If $w \equiv (11)^p(2221111)(222112221111)^r(222)^q \equiv 1^{2p}\ 1\ (21)^3\ (21)^{6r}\ 2^{3q} \in \text{IRR}(R)$.

If $w \equiv (11)^p(2221111)(22211)^{s_2}(222)^q \equiv 1^{2p}\ 1\ (21)^3\ (21)^{2s_2}\ 2^{s_2}\ 2^{3q} \in \text{IRR}(R)$.

If

$w \equiv (11)^p(222112221111)^r(22211)^{s_2}(222)^q \equiv 1^{2p}\ (21)^{6r}\ (21)^{2s_2}\ 2^{s_2}\ 2^{3q} \in \text{IRR}(R)$.

If $w \equiv (11)^p(2221111)(222112221111)^r(22211)^{s_2}(222)^q$

$\equiv 1^{2p}\ 1\ (21)^3\ (21)^{6r}\ (21)^{2s_2}\ 2^{s_2}\ 2^{3q} \in \text{IRR}(R)$.

Now we can complete the proof and determine w', as follows:

$$\begin{aligned} w' &\equiv \left(11\right)^p\ \left(\ (1)(21)^3\ \right)^{s_3}\ \left(\ (21)^6\ \right)^r\ \left(\ (21)^{2s_2}(2)^{s_2}\ \right)\ \left(222\right)^q \\ &\equiv 1^{(2p+s_3)}\ (21)^{(3s_3+6r+2s_2)}\ 2^{(3q+s_2)}. \end{aligned}$$

where w' is in $\text{IRR}(R)$ and in normal form for P_2, by Corollary 6.4.1. □

Lemma 6.7.11. *The mapping $\psi : M \to F_1$ is a well-defined bijection.*

Proof. By Lemma 6.7.9 the mapping ψ is a well-defined epimorphism that is a surjective homomorphism. Therefore it remains to prove that ψ is an injective mapping. The proof will be by contradiction and we will assume that ψ is not injective.

Let $t_1, t_2 \in B^*$ and our assumption (for a contradiction) is that $t_1/\rho \neq t_2/\rho$ in the monoid M but $(t_1/\rho)\psi = (t_2/\rho)\psi$ in the submonoid F_1. Next, without loss of generality, let t_1 and t_2 be in normal form as any word in B^* can be reduced to a normal form by Lemma 6.7.6. Let

$$t_1 \equiv (a)^p(baa)^{s_3}(babaa)^r(ba)^{s_2}(b)^q \text{ and } t_2 \equiv (a)^{p'}(baa)^{s'_3}(babaa)^{r'}(ba)^{s'_2}(b)^{q'}$$

with integers $p, p', r, r', q, q' \geq 0$; s_3 and $s'_3 \in \{0, 1\}$; s_2 and $s'_2 \in \{0, 1, 2\}$.

Let $(t_1/\rho)\psi = t_1\phi = w_1/\eta$ and $w_1 \xrightarrow{*}_R w_1' \in \mathrm{IRR}(R)$. Similarly we will let $(t_2/\rho)\psi = t_2\phi = w_2/\eta$ and $w_2 \xrightarrow{*}_R w_2' \in \mathrm{IRR}(R)$. By our assumption, let $w_1'/\eta = w_2'/\eta$. As w_1' and w_2' are in normal form, this implies that $w_1' \equiv w_2'$.

By Lemma 6.7.10 we have

$$w_1' \equiv 1^{(2p+s_3)}\,(21)^{(3s_3+6r+2s_2)}\,2^{(3q+s_2)} \text{ and}$$
$$w_2' \equiv 1^{(2p'+s_3')}\,(21)^{(3s_3'+6r'+2s_2')}\,2^{(3q'+s_2')}.$$

By definition s_3 and s_3' are in $\{0,1\}$ and therefore we must have $p = p'$ and also $s_3 = s_3'$. Similarly, by definition s_2 and s_2' are in $\{0,1,2\}$ and therefore we must have $q = q'$ and also $s_2 = s_2'$. Finally, as $s_3 = s_3'$ and $s_2 = s_2'$ then we must have $r = r'$. This means that $t_1 \equiv t_2$ and $t_1/\rho = t_2/\rho$ which contradicts our assumption and so ψ is injective, which completes our proof. □

6.7.3 Proof of new theorem

We begin this section by returning to our original proposition and we now have everything in place to prove it as a new theorem.

Theorem 6.7.12. *Let $A = \{1,2\}$ and $R = \{(221,212),(211,121)\}$. Then the plactic monoid P_2 is defined by the presentation $\langle A \mid R\rangle$.*

Let F_1 be the submonoid of P_2 generated by $Y_1 = \{1^2, 2^3\}$. Set $B = \{a,b\}$ and let Q be the subset of $B^* \times B^*$ consisting of all the following pairs:
(i) $(baaa, abaa)$,
(ii) (bba, bab),
(iii) $(baabaa, ababaa)$,
(iv) $(bababa, babaab)$.
Then (B,Q) is a finite complete rewriting system defining the monoid F_1 where a and b correspond to the generators 1^2 and 2^3, respectively.

Proof. In Lemma 6.7.5 we proved that the presentation $\langle B \mid Q\rangle$ defining the monoid M is a FCRS. In Definitions 6.7.7 and Lemma 6.7.9 we have defined mappings $\phi : B^* \to F_1$ and $\psi : M \to F_1$.

The presentation $\langle B \mid Q\rangle$ is generated by the classes b/ρ where $b \in B$ and ψ is a one to one mapping to the classes $(b/\rho)\psi = y/\eta$ where $y \in Y_1$, which generate the submonoid F_1. By Lemma 6.7.8 all the relations in Q also hold in F_1. Also, in Lemma 6.7.11 we prove that ψ is a well-founded bijective homomorphism, in other words an isomorphism between M and F_1.

As the presentation $\langle B \mid Q\rangle$ is a FCRS which defines the monoid M and we have an isomorphism between M and F_1, then the presentation $\langle B \mid Q\rangle$ is a finite complete rewriting system which defines the monoid F_1. □

6.8 Submonoids of the plactic monoid P_2 generated by the set $Y = \{1^i, 2^j\}$

In previous subsections we have considered the submonoids generated by:
$X = \{1, 2^i\}$, $X' = \{1^i, 2\}$, $Y_1 = \{1^2, 2^3\}$,
and proved that they have finite presentations which are also FCRS.

In this subsection we consider the same questions for the generating set $Y = \{1^i, 2^j\}$ where $i, j \in \mathbb{N}$ and $i < j$. In order to find a presentation, the first step is to look for a normal form, together with a set of rewrite rules. In the remainder of this subsection results are presented for specific generating sets and a conjecture for the generating set Y. First some specific examples.

Conjecture 6.8.1. *Let $A = \{1, 2\}$ and $R = \{(221, 212), (211, 121)\}$. Then the plactic monoid P_2 is defined by the presentation $\langle A \mid R\rangle$.*

Let F_2 be the submonoid of P_2 generated by $Y_2 = \{1^3, 2^8\}$. Set $B = \{a, b\}$ and let Q be the subset of $B^* \times B^*$ consisting of all the following pairs:
(0) (bba, bab),

(1) $(baaaa, abaaa)$,

(2) $(baaabaaabaaa, abaabaaabaaa)$,

(3) $(baba, baab)$,

(4) $(baabaaba, baabaaab)$,

(5) $(baabaaabaaba, baabaaabaaab)$.

Then (B, Q) *is a finite complete rewriting system defining the monoid* F_2 *where* a *and* b *correspond to the generators* 1^3 *and* 2^8*, respectively. Moreover, the irreducible words with respect to the rewriting system* (B, Q) *are precisely the words of the form:*

$$(a)^p\ (baaa)^{s_8}\ (baabaaabaaa)^r\ (baabaaa)^{s_{15}}\ (baa)^{s_6}\ (ba)^{s_3}\ (b)^q$$

where $p, q, r \geq 0$, $s_8 \in \{0, 1, 2\}$, $s_{15} \in \{0, 1\}$, $s_6 \in \{0, 1, 2\}$, $s_3 \in \{0, 1\}$ *and not* $s_6 = 2$ *with* $s_3 = 1$ *and not* $s_{15} = 1$ *with* $s_6 = 2$*. Hence this set of words gives a set of normal forms for the elements of the monoid* F_2 *defined by the presentation* $\langle B \mid Q\rangle$.

Conjecture 6.8.2. *Let* $A = \{1, 2\}$ *and* $R = \{(221, 212), (211, 121)\}$*. Then the plactic monoid* P_2 *is defined by the presentation* $\langle A \mid R\rangle$.

Let F_3 *be the submonoid of* P_2 *generated by* $Y_3 = \{1^5, 2^{12}\}$*. Then* F_3 *admits a presentation by a finite complete rewriting system.*

Let $B = \{a, b\}$ *be an alphabet where* a *and* b *correspond to the generators* 1^3 *and* 2^8*, respectively. Then the set of normal forms for the elements of the monoid* F_3 *are precisely the words of the form:*

$$(a)^p\ (baaa)^{s_{12}}\ (baabaaa)^{s_{24}}\ (baabaabaaabaabaaa)^r$$

$$(baabaabaaa)^{s_{35}}\ (baa)^{s_{10}}\ (ba)^{s_5}\ (b)^q$$

where $p, q, r \geq 0$, $s_{12} \in \{0, 1\}$, $s_{24} \in \{0, 1, 2\}$, $s_{35} \in \{0, 1\}$, $s_{10} \in \{0, 1, 2, 3\}$, $s_5 \in \{0, 1\}$*. There would also be some combinations of* s_n*'s which would not be* *possible.*

Conjecture 6.8.3. *Let P_2 be the plactic monoid of rank 2 generated by $\{1, 2\}$. For $i, j \in \mathbb{N}$ let $M_{i,j}$ be the submonoid of P_2 generated by $Y = \{1^i, 2^j\}$. Then $M_{i,j}$ admits a presentation by a finite complete rewriting system.*

Currently this is just a conjecture. A discussion of how we might approach proving this conjecture is given below. By Lemma 6.2.6 and by symmetry, it suffices to prove the result for the case $i < j$ with i and j coprime.

Let $B = \{a, b\}$ be an alphabet where a and b correspond to the generators 1^i and 2^j respectively. Then the conjecture is that the normal forms for the elements of the submonoid $M_{i,j}$ will have the following form:

$(a)^p$ [left mixed terms](lowest common multiple term)^r[right mixed terms]$(b)^q$

where $p, q, r \geq 0$. The lowest common multiple term will always be present and consists of a single word which is a combination of the generators a and b. There will be j occurrences of the letter a and i occurrences of the letter b and it will look something like this, $\big(\ b\ (a)^{x_1}\ \ b\ (a)^{x_2}\ \dots\ \ b\ (a)^{x_{(i-1)}}\ \ b(a)^{x_i}\ \big)$ with $x_1, x_2, \dots, x_i \in \mathbb{N}$. This term can be identified as the term which is raised to the power r in Conjectures 6.8.1 and 6.8.2. Looking at these examples it is clear that the number of terms to the left and the right of this term can vary depending on the relative values of i and j. As such it may be possible to find a method by which to determine the sets of left mixed terms and right mixed terms.

We conjecture that a set of rewrite rules can be derived from the normal form. It may not be possible to determine a generalised form for these, rather it will be possible to find a method by which to determine them.

6.9 Potential future work

6.9.1 Submonoids of P_2, P_3 or P_n

(i) Develop conjecture 6.8.3 and see if it is possible to define a method for finding a presentation which is a finite complete rewriting system.

(ii) Extend current results to the plactic monoid P_3 where the generating set for the submonoid of P_3 has three variables e.g. $X = \{1^i, 2^j, 3^k\}$. Again we look to find a presentation for the submonoid which is a finite complete rewriting system.

(iii) Further generalise the generating set to define different submonoids of the plactic monoid P_3, for example:

$$X = \{1^{i_1}, 1^{i_2}, \ldots, 1^{i_r}, 2^{j_1}, 2^{j_2}, \ldots, 2^{j_s}, 3^{k_1}, 3^{k_2}, \ldots, 3^{k_t}\}.$$

(iv) Let P_n be the plactic monoid generated by the set $A = \{1, 2, \ldots, n\}$. Let M be the submonoid of P_n generated by the set $X = \{w_1, w_2\}$ where $w_1, w_2 \in A^*$. Is the submonoid M finitely presented and does it admit a presentation by a finite complete rewriting system?

6.9.2 Open questions

(i) Is it possible to classify all finitely generated submonoids of the plactic monoid P_n for various values of $n \in \mathbb{N}$?

(ii) Are all finitely generated submonoids of the plactic monoid P_n finitely presented, have finite derivation type or admit a presentation by a finite complete rewriting system?

Bibliography

[1] H. Ayik and N. Ruškuc. Generators and relations of Rees matrix semigroups. *Proceedings of the Edinburgh Mathematical Society*, 42(3):481–495, 10 1999.

[2] F. Baader and T. Nipkow. *Term rewriting and all that*. Cambridge university press, 1999.

[3] R.V. Book and F. Otto. *String-rewriting systems*. Springer, 1993.

[4] A.J. Cain, R.D. Gray, and A. Malheiro. Finite Gröbner–Shirshov bases for plactic algebras and biautomatic structures for plactic monoids. *Journal of Algebra*, 423(Supplement C):37–53, 2015.

[5] C.M. Campbell, E.F. Robertson, N. Ruškuc, and R.M. Thomas. Reidemeister - Schreier type rewriting for semigroups. *Semigroup Forum*, 51:47–62, 1995.

[6] C.M. Campbell, E.F. Robertson, N. Ruškuc, and R.M. Thomas. On subsemigroups of finitely presented semigroups. *Journal of Algebra*, 180(1):1–21, 1996.

[7] C.M. Campbell, E.F. Robertson, N. Ruškuc, and R.M. Thomas. Presentations for subsemigroups — applications to ideals of semigroups. *Journal of Pure and Applied Algebra*, 124(1):47–64, 1998.

[8] F. Cedó and J. Okniński. Plactic algebras. *Journal of Algebra*, 274(1):97–117, 2004.

[9] A. Clayton and N. Ruškuc. On the number of subsemigroups of direct products involving the free monogenic semigroup. *J. Aust. Math. Soc.*, 109(1):24–35, 2020.

[10] 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.

[11] A.H. Clifford and G.B. Preston. *The algebraic theory of semigroups. Vol. II.* Mathematical Surveys, No. 7. American Mathematical Society, Providence, R.I., 1967.

[12] D.E. Cohen. String rewriting and homology of monoids. *Mathematical. Structures in Comp. Sci.*, 7(3):207–240, June 1997.

[13] L. Descalço and N. Ruškuc. Subsemigroups of the bicyclic monoid. *International Journal of Algebra and Computation*, 15(01):37–57, 2005.

[14] L. Descalço and N. Ruškuc. Properties of the subsemigroups of the bicyclic monoid. *Czechoslovak Mathematical Journal*, 58(2):311–330, 2008.

[15] W. Fulton. *Young tableaux: with applications to representation theory and geometry*, volume 35. Cambridge University Press, 1997.

[16] R.D. Gray. *A graph theoretic approach to combinatorial problems in semigroup theory.* PhD thesis, University of St Andrews, 2006.

[17] R.D. Gray and A. Malheiro. A monoid is defined by a complete semigroup presentation if and only if it is defined by a complete monoid presentation. Technical report, Technical Report, Centro de Álgebra da Universidade de Lisboa, 2010 (http://pessoa. fct. unl. pt/ajm/TechRep/CRSforMonoids. pdf), 2010.

[18] R.D. Gray and A. Malheiro. Finite complete rewriting systems for regular semigroups. *Theoretical Computer Science*, 412(8):654–661, 2011.

[19] R.D. Gray and A. Malheiro. Homotopy bases and finite derivation type for subgroups of monoids. *Journal of Algebra*, 410:53–84, 2014.

[20] R.D. Gray, A. Malheiro, and S.J. Pride. Homotopy bases and finite derivation type for Schützenberger groups of monoids. *Journal of Symbolic Computation*, 50:50–78, 2013.

[21] R.D. Gray and N. Ruškuc. Generators and relations for subsemigroups via boundaries in Cayley graphs. *Journal of Pure and Applied Algebra*, 215(11):2761–2779, 2011.

[22] J. A. Green. On the structure of semigroups. *Ann. of Math. (2)*, 54:163–172, 1951.

[23] P.A. Grillet. *Semigroups: an introduction to the structure theory*, volume 193. CRC Press, 1995.

[24] D.F. Holt, B. Eick, and E.A. O'Brien. *Handbook of computational group theory*. CRC Press, 2005.

[25] J.M. Howie. *Fundamentals of semigroup theory*. Oxford University Press, 1996.

[26] J.M. Howie and N. Ruškuc. Constructions and presentations for monoids. *Communications in Algebra*, 22(15):6209–6224, 1994.

[27] D. Kapur and P. Narendran. A finite thue system with decidable word problem and without equivalent finite canonical system. *Theoretical Computer Science*, 35:337–344, 1985.

[28] D. Knuth. Permutations, matrices, and generalized young tableaux. *Pacific journal of mathematics*, 34(3):709–727, 1970.

[29] Y. Kobayashi. Finite homotopy bases of one-relator monoids. *Journal of Algebra*, 229(2):547–569, 2000.

[30] Ł. Kubat and J. Okniński. Gröbner-Shirshov bases for plactic algebras. *Algebra Colloq.*, 21(4):591–596, 2014.

[31] A. Lascoux, B. Leclerc, and J-Y. Thibon. The plactic monoid. In *Algebraic Combinatoric on Words*, page 10pp. Cambridge University Press, 2002.

[32] A. Lascoux and M-P. Schutzenberger. Le monoide plaxique,(proc. colloqu. naples 1978). *Quad. Ricerca Sci*, 109, 1981.

[33] B. Leclerc and J-Y. Thibon. The Robinson-Schensted correspondence, crystal bases, and the quantum straightening at $q = 0$. *Electronic Journal of Combinatorics*, 3(2):Research paper 11, 1996.

[34] M. Lothaire. *Algebraic combinatorics on words*, volume 90 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2002.

[35] R.C. Lyndon and P.E. Schupp. *Combinatorial group theory*. Springer, 2015.

[36] W. Magnus, A. Karrass, and D. Solitar. *Combinatorial group theory: Presentations of groups in terms of generators and relations*. Courier Corporation, 2004.

[37] A. Malheiro. Complete rewriting systems for codified submonoids. *International Journal of Algebra and Computation*, 15(02):207–216, 2005.

[38] A. Malheiro. Finite derivation type for Rees matrix semigroups. *Theoretical Computer Science*, 355(3):274–290, 2006.

[39] A. Malheiro. On trivializers and subsemigroups. In *Semigroups and formal languages*, pages 188–204. World Scientific, 2007.

[40] A. Malheiro. Finite derivation type for large ideals. *Semigroup Forum*, 78(3):450, 2008.

[41] J. Meakin. The Rees construction in regular semigroups. In *Semigroups (Szeged, 1981)*, volume 39 of *Colloq. Math. Soc. János Bolyai*, pages 115–155. North-Holland, Amsterdam, 1985.

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

[43] F. Otto and Y. Kobayashi. Properties of monoids that are presented by finite convergent string-rewriting systems—a survey. In *Advances in algorithms, languages, and complexity*, pages 225–266. Kluwer Acad. Publ., Dordrecht, 1997.

[44] S.J. Pride and J. Wang. Subgroups of finite index in groups with finite complete rewriting systems. *Proceedings of the Edinburgh Mathematical Society*, 43(1):177–183, 2000.

[45] D. Rees. On semi-groups. *Proc. Cambridge Philos. Soc.*, 36:387–400, 1940.

[46] N. Ruškuc. Presentations for subgroups of monoids. *Journal of Algebra*, 220(1):365–380, 1999.

[47] N. Ruškuc. *Semigroup presentations.* PhD thesis, University of St Andrews, 1995.

[48] N. Ruškuc. On large subsemigroups and finiteness conditions of semigroups. *Proceedings of the London Mathematical Society*, 76(2):383–405, 003 1998.

[49] C. Schensted. Longest increasing and decreasing subsequences. *Canadian Journal of Mathematics*, 13:179–191, 1961.

[50] W.Y. Sit and M-K. Siu. On the subsemigroups of N. *Mathematics Magazine*, 48(4):225–227, 1975.

[51] C.C. Squier, F. Otto, and Y. Kobayashi. A finiteness condition for rewriting systems. *Theoretical Computer Science*, 131(2):271–294, 1994.

[52] A. Suschkewitsch. Über die endlichen Gruppen ohne das Gesetz der eindeutigen Umkehrbarkeit. *Mathematische Annalen*, 99(1):30–50, 1928.

[53] J. Wang. Finite complete rewriting systems and finite derivation type for small extensions of monoids. *Journal of Algebra*, 204(2):493–503, 1998.

[54] K.B. Wong and P.C. Wong. On finite complete rewriting systems and large subsemigroups. *Journal of Algebra*, 345(1):242–256, 2011.

[55] L. Zhang. Applying rewriting methods to special monoids. *Math. Proc. Cambridge Philos. Soc.*, 112(3):495–505, 1992.