A graph theoretic approach to combinatorial problems in semigroup theory

Robert Gray

School of Mathematics and Statistics
University of St Andrews

Thesis submitted for the Degree of Doctor of Philosophy
University of St Andrews

· November 5, 2005 ·

Abstract

The use of graph theory has become widespread in the algebraic theory of semigroups. In this context, the graph is mainly used as a visual aid to make presentation clearer and the problems more manageable. Central to such approaches is the Cayley graph of a semigroup. There are also many variations on the idea of the Cayley graph, usually special kinds of subgraph or factor graph, that have become important in their own right. Examples include Schützenberger graphs, Schreier coset graphs and Van Kampen diagrams (for groups), Munn trees, Adian graphs, Squier complexes, semigroup diagrams, and graphs of completely 0-simple semigroups. Also, the representation of elements in finite transformation semigroups as digraphs has proved a useful tool.

This thesis consists of several problems in the theory of semigroups with the common feature that they are all best attacked using graph theory. The thesis has two parts. In the first part combinatorial questions for finite semigroups and monoids are considered. In particular, we look at the problem of finding minimal generating sets for various endomorphism monoids and their ideals. This is achieved by detailed analysis of the generating sets of completely 0-simple semigroups. This investigation is carried out using a bipartite graph representation.

The second part of the thesis is about infinite semigroup theory, and in particular some problems in the theory of semigroup presentations. In particular we consider the general problem of finding presentations for subsemigroups of finitely presented semigroups. Sufficient conditions are introduced that force such a subsemigroup to be finitely presented. These conditions are given in terms of the position of the subsemigroup in the parent semigroup's left and right Cayley graphs.

Declarations

I, Robert Gray, hereby certify that this thesis, which is approximately 60,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Signed Date 14/02/06

I was admitted as a research student in September 2002 and as a candidate for the degree of PhD in September 2003; the higher study for which this is a record was carried out in the University of St Andrews between 2002 and 2005.

Signed Date 14/02/06

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

Signed Date 14/02/06

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Signature of Supervisor Date 14/02/06

Acknowledgments

I would like to take this chance to thank those people who have made this thesis possible.

Firstly I would like to thank my family for the constant support and encouragement they have provided over the years. I would especially like to thank my mother, father and my two sisters Heather and Mary.

Next I would like to thank my supervisor Professor Nik Ruškuc for his encouragement and guidance during my PhD. I would also like to thank the rest of the staff at the university for all the help that they have provided.

Finally I would like to thank all my friends and colleagues who have made the past three years so enjoyable. I would especially like to thank Catarina, for proof reading my papers and thesis, Lizzie, for being a wonderful office mate, and Peter, for the interesting maths conversations at the breakfast table.

Contents

List of Figures

Preface

This thesis is a collection of problems, and some solutions, from the theory of semigroups. It is divided into two parts. In the first part finite semigroups are considered and some combinatorial questions are investigated. In the second part infinite semigroups are considered, in particular we look at a number of questions in the theory of semigroup presentations. A summary of the topics covered in this thesis is given below.

In Chapter 1 the necessary semigroup theory and graph theory preliminaries are given. In Chapter 2 minimal generating sets for finite completely 0-simple semigroups are investigated. In particular, we give a formula for the minimum cardinality of a generating set for such a semigroup. Several applications to various finite semigroups of transformations are also given. In Chapter 3 semibands (idempotent generated regular semigroups) are considered. Connections between minimal generating sets of idempotents of semibands, and matchings in bipartite graphs are explored. Necessary and sufficient conditions for a completely 0-simple semigroup to have an extremal idempotent generating set are given. These results are applied to ideals of the full transformation and the general linear semigroups. A consequence of this is a result giving necessary and sufficient conditions for a subset of a two-sided ideal of the general linear semigroup to be a minimal generating set.

Independence algebras and their endomorphism monoids are the subject of Chapters 4 and 5. The main results of these chapters generalise combinatorial results of Howie and McFadden for the full transformation semigroup, and results of Dawlings for the general linear semigroup, to the more general context of endomorphism monoids of independence algebras. The necessary and sufficient conditions for completely 0-simple semigroups to have extremal idempotent generating sets, established in Chapter 3, provide the basis on which the main results of Chapters 4 and 5 are constructed. We finish our consideration of finite semigroups in Chapter 6 where an extremal problem for subsemigroups of the full transformation semigroup is discussed. The largest order completely simple subsemigroups of the full transformation semigroup are described using an argument that involves counting the number of distinct r-colourings of members of a certain family of r-partite graphs.

Chapter 7 is the place that we begin our study of infinite semigroups and in particular the theory of semigroup presentations. Given a finitely generated semigroup S and subsemigroup T of S we define the notion of the boundary of

T in S which, intuitively, describes the position of T inside the left and right Cayley graphs of S. We prove that if S is finitely generated and T has a finite boundary in S then T is finitely generated. We also prove that if S is finitely presented and T has a finite boundary in S then T is finitely presented. Several corollaries and examples are given. In Chapter 8 we continue with the subject-matter introduced in Chapter 7. In particular, the boundaries of left and right unitary subsemigroups are analysed which leads to results for presentations of unitary subsemigroups. Several applications are given.

The main results of the thesis have been written up as a series of five research articles; see [42], [43], [44], [45], [46]. For all of the main results of this thesis, approaching the problem from the point of view of graphs proved to be the "right way" of thinking about the problem. I hope the ideas that appear here will be of interest to those who read them and especially to others who, like me, think in pictures.

Chapter 1

Preliminaries

1.1 Semigroup theory preliminaries

In this section all of the basic semigroup theory needed to understand the results of the thesis will be presented. All of the definitions and results are standard and can be found in any introductory text on the subject (see for example [57], [64], [49], [52] or [21]).

Subsemigroups and generating sets

A semigroup is a pair $(S, \cdot)$ where S is a non-empty set and $\cdot$ is a binary operation defined on S that satisfies the *associative* law

$$(x \cdot y) \cdot z = x \cdot (y \cdot z)$$

for all $x, y, z \in S$. The product of two elements x and y is usually written just as xy rather than $x \cdot y$. If a semigroup contains an element 1 with the property that $x1 = 1x = x$ for all $x \in S$ then we call 1 the *identity element* of the S and we call S a *monoid*. If a semigroup contains an element 0 that satisfies $x0 = 0x = 0$ for all $x \in S$ then 0 is called a *zero element* of the semigroup. A semigroup can have at most one identity element and at most one zero element.

We use S^1 and S^0 to denote the semigroup S with an identity or a zero adjoined, respectively. That is,

$$S^1 = \begin{cases} S & \text{if } S \text{ has an identity element} \\ S \cup \{1\} & \text{otherwise} \end{cases}$$

and

$$S^0 = \begin{cases} S & \text{if } S \text{ has a zero element} \\ S \cup \{0\} & \text{otherwise.} \end{cases}$$

If A and B are subsets of a semigroup we define the product of the two sets as $AB = \{ab : a \in A, b \in B\}$. In the special case of singleton subsets $A = \{a\}$ we write aB rather than $\{a\}B$. So for example, $S^1 a = Sa \cup \{a\}$.

A non-empty subset T of S is called a *subsemigroup* if it is closed under multiplication. Let S be a semigroup and let $\{T_i : i \in I\}$ be an indexed set of subsemigroups of S. Then if the set $\bigcap_{i \in I} T_i$ is non-empty, it is a subsemigroup of S. In particular, for any non-empty subset A of S the intersection of all the subsemigroups of S that contain A is non-empty and is a subsemigroup of S. We use $\langle A \rangle$ to denote this subsemigroup and call it the *subsemigroup of S generated* *by the set* A. The subsemigroup $\langle A \rangle$ is the set of all elements in S that can be written as a finite product of elements of A.

An element $e \in S$ is called an *idempotent* if it satisfies $e^2 = e$. We use $E(S)$ and $F(S)$ to denote the idempotents of S and the subsemigroup generated by the set of idempotents, respectively. A *band* is a semigroup such that every element is an idempotent.

Homomorphisms and congruences

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

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

An equivalence relation which is stable under left multiplication is called a *left congruence* and a relation that is both a left and right congruence is called a *(two-sided) congruence* on S. If ρ is a congruence on S then we can define a binary operation on the quotient set S/ρ by

$$(a\rho)(b\rho) = (ab)\rho.$$

A map $\phi : S \to T$ where S and T are semigroups is called a *homomorphism* if for all $x, y \in S$

$$(xy)\phi = x\phi y\phi.$$

If S and T are monoids then, to be called a monoid homomorphism, ϕ must also satisfy $1_S\phi = 1_T$. A homomorphism that is injective will be called a *monomorphism* and if it is surjective it will be called an *epimorphism.* Also, a homomorphism is called an *isomorphism* if it is bijective. When there exists an epimorphism from S onto T we say that T is a *homomorphic image* of S. If there is an isomorphism $\phi : S \to T$ we say that S and T are isomorphic and write $S \cong T$. A homomorphism from S to itself is called an *endomorphism* and an isomorphism from S to itself is called an *automorphism.* The set of all endomorphisms of S, under composition of maps, forms a monoid. We call this monoid the *endomorphism monoid* of S and denote it by $\text{End}(S)$. Similarly, the set of automorphisms forms a group that is denoted $\text{Aut}(S)$ and is called the *automorphism group* of S. Given a map $\phi : S \to T$ we define

$$\ker \phi = \{(x,y) \in S \times S : x\phi = y\phi\}$$

and call this the *kernel* of the map ϕ. The first isomorphism theorem for semigroups tells us that with every epimorphism $\phi : S \to T$ the kernel $\ker\phi$ is a congruence on S and $S/\ker\phi \cong T$. Conversely, if ρ is a congruence on S then the map $\phi : S \to S/\rho$ defined by $x\phi = x/\rho$ is an epimorphism from S onto the factor semigroup S/ρ.

Ideals and Rees quotients

A subsemigroup T of a semigroup S that satisfies $TS \subseteq T$ is called a *right ideal*. Dually, if $ST \subseteq T$ then T is called a *left ideal* and T is called a (two-sided) ideal if it is both a left and a right ideal. An ideal I of S is called *proper* if $I \neq S$. If I is a proper ideal of a semigroup S then

$$\rho_I = \{(s,s) : s \in S\} \cup (I \times I)$$

is a congruence on S. It is useful to think of S/ρ_I as $(S \setminus I) \cup \{0\}$ where all products not falling in $S \setminus I$ are equal to zero. We shall call a congruence of this type a *Rees congruence*, and if a homomorphism $\phi : S \to T$ is such that $\ker\phi$ is a Rees congruence we shall say that ϕ is a *Rees homomorphism*. We shall normally write S/I rather than S/ρ_I and call this the *Rees quotient of S with respect to I*.

Regular semigroups, Green's relations and the structure of a $\mathcal{D}$-class

Green's relation were first introduced in [48]. They describe the ideal structure of a semigroup. Since their introduction they have played a central role in the structure theory of semigroups. We now define Green's relations $\mathcal{R}$, $\mathcal{L}$, $\mathcal{J}$, $\mathcal{D}$ and $\mathcal{H}$ and give some of their basic properties.

Let S be a semigroup and let $s \in S$. The *principal right, left and two-sided ideals* generated by s are the sets $sS^1 = sS \cup \{s\}$, $S^1s = Ss \cup \{s\}$ and $S^1sS^1 = sS \cup Ss \cup SsS \cup \{s\}$, respectively. For $s,t \in S$ we say that s and t are $\mathcal{R}$-related, writing $s\mathcal{R}t$, if s and t generate the same principal right ideal. We say they are $\mathcal{L}$-related if they generate the same principal left ideal, in which case we write $s\mathcal{L}t$. Also, we say s and t are $\mathcal{J}$-related, writing $s\mathcal{J}t$, if they generate the same principal two-sided ideal. We define $\mathcal{H} = \mathcal{R} \cap \mathcal{L}$ and $\mathcal{D} = \mathcal{R} \circ \mathcal{L} = \mathcal{L} \circ \mathcal{R}$: the composition of the binary relations $\mathcal{R}$ and $\mathcal{L}$. Each of these relations is an equivalence relation on the semigroup S and we call the corresponding equivalence classes the $\mathcal{R}$-, $\mathcal{L}$-, $\mathcal{J}$-, $\mathcal{D}$- and $\mathcal{H}$-classes, respectively of S.

Given an element $s \in S$, we will use R_s, L_s, J_s, D_s and H_s to denote the $\mathcal{R}$-,

$\mathcal{L}$-, $\mathcal{J}$-, $\mathcal{D}$- and $\mathcal{H}$-classes, respectively, of s in S. Since $\mathcal{R}$, $\mathcal{L}$ and $\mathcal{J}$ are defined in terms of ideals, ordering the right, left and two-sided ideals of S by inclusion induces a partial order on these equivalence classes given by

$$\begin{aligned} L_a \leq L_b &\Leftrightarrow S^1a \subseteq S^1b \\ R_a \leq R_b &\Leftrightarrow aS^1 \subseteq bS^1 \\ J_a \leq J_b &\Leftrightarrow S^1aS^1 \subseteq S^1bS^1. \end{aligned}$$

In finite semigroups the relations $\mathcal{D}$ and $\mathcal{J}$ coincide and, when working with finite semigroups, we write $D_s \leq D_t$ to mean $J_s \leq J_t$.

Each $\mathcal{D}$-class of a semigroup S is a union of $\mathcal{R}$-classes and also is a union of $\mathcal{L}$-classes. Moreover, $a\mathcal{D}b$ if and only if $R_a \cap L_b \neq \varnothing$ which is true if and only if $R_b \cap L_a \neq \varnothing$. If D is a $\mathcal{D}$-class of S and if $a, b \in D$ are $\mathcal{R}$-related in S, say with $as = b$ and $bs' = a$, then the right translation $\rho_s : S \to S$ defined by $x\rho_s = xs$ maps L_a to L_b. The map $\rho_{s'} : S \to S$ maps L_b back to L_a and the composition of the maps $\rho_s\rho_{s'} : S \to S$ acts as the identity map on L_a. Moreover, the map ρ_s is $\mathcal{R}$-class preserving in the sense that it maps each $\mathcal{H}$-class on L_a in a 1-1 manner onto the corresponding ($\mathcal{R}$-equivalent) $\mathcal{H}$-class of L_b. There is a dual result for $\mathcal{L}$-classes. These results, collectively, are known as Green's lemma.

It is often useful to visualise a $\mathcal{D}$-class of a semigroup using a so called *egg-box diagram.* An egg-box diagram of a $\mathcal{D}$-class D is a grid whose rows represent the $\mathcal{R}$-classes of D, its columns represent the $\mathcal{L}$-classes of D, and the intersections of the rows and columns, that is, the cells of the grid, represent the $\mathcal{H}$-classes of the semigroup. Egg-box diagrams will be found scattered amongst the chapters of this thesis. They provide a useful tool for visualising semigroups.

The $\mathcal{H}$-classes of a given $\mathcal{D}$-class all have the same size. Each $\mathcal{H}$-class H of S is either a subgroup of S or satisfies $H^2 \cap H = \varnothing$. An $\mathcal{H}$-class is a subgroup of S if and only if it contains an idempotent (which will act as the identity of that subgroup). We call the $\mathcal{H}$-classes that contain idempotents the *group $\mathcal{H}$-classes* of S. Any two group $\mathcal{H}$-classes in a given $\mathcal{D}$-class are isomorphic.

An element $a \in S$ is called *regular* if there exists $x \in S$ such that $axa = a$. The semigroup S is said to be regular if all of its elements are regular. If D is a $\mathcal{D}$-class then either every element in D is regular or none of them are. The $\mathcal{D}$-classes that have regular elements are called the *regular $\mathcal{D}$-classes.* In a regular $\mathcal{D}$-class each $\mathcal{R}$-class and each $\mathcal{L}$-class contains an idempotent. If $a \in S$ then we say that a' is an inverse of a if

$$aa'a = a, \quad a'aa' = a'.$$

An element has an inverse if and only if that element is regular. The following lemma is used extensively in the thesis.

Lemma 1.1. *Let a, b be elements in a $\mathcal{D}$-class D. Then $ab \in R_a \cap L_b$ if and only if $L_a \cap R_b$ contains an idempotent.*

0-simple semigroups, the Rees theorem and principal factors

A semigroup is called simple if it has no proper ideals. This is equivalent to saying that the semigroup has a single $\mathcal{J}$-class. A completely simple semigroup is a simple semigroup that has minimal left and right ideals. Every finite simple semigroup is completely simple.

A left zero semigroup is a semigroup in which every element acts as a left zero (i.e. in which $xy = x$ for all $x, y \in S$). A right zero semigroup is one in which every element acts as a right zero. Note that left and right zero semigroups are just special kinds of completely simple semigroup.

A semigroup is called 0-simple if $\{0\}$ and S are its only ideals (and $S^2 \neq \{0\}$). This is equivalent to saying that $\{0\}$ and $S \setminus \{0\}$ are its only $\mathcal{J}$-classes (and $S^2 \neq \{0\}$). A semigroup is 0-simple if and only if $SaS = S$ for every $a \neq 0$ in S. A semigroup S is said to be *completely 0-simple* if it is 0-simple and has 0-minimal left and right ideals. By a 0-minimal left (respectively right) ideal we mean a left (respectively right) ideal that is minimal within the set of all non-zero left (respectively right) ideals ordered by inclusion. Every finite 0-simple semigroup is completely 0-simple.

0-simple semigroups occur naturally 'inside' arbitrary semigroups appearing as *principal factors* of $\mathcal{J}$-classes. Let J be some $\mathcal{J}$-class of a semigroup S. Then the principal factor of S corresponding to J is the set $J^* = J \cup \{0\}$ with multiplication

$$s * t = \begin{cases} st & \text{if } s, t, st \in J \\ 0 & \text{otherwise.} \end{cases}$$

The semigroup J^* is either a semigroup with zero multiplication or is a 0-simple semigroup. The following construction, due to Rees, gives a method for building completely 0-simple semigroups

Let G be a group, let I, Λ be non-empty index sets and $P = (p_{\lambda i})$ a *regular* $\Lambda \times I$ matrix over $G \cup \{0\}$ (where regular means that every row and every column of P has at least one non-zero entry). Then $S = \mathcal{M}^0[G; I, \Lambda; P]$, *the* $I \times \Lambda$ *Rees matrix semigroup over the* 0*-group* $G \cup \{0\}$ *with sandwich matrix* P, is the semigroup $(I \times G \times \Lambda) \cup \{0\}$ with multiplication defined by

$$(i, g, \lambda)(j, h, \mu) = \begin{cases} (i, gp_{\lambda j}h, \mu) & \text{if } p_{\lambda j} \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

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

The semigroup $S = \mathcal{M}^0[G; I, \Lambda; P]$ is completely 0-simple. Moreover, by [57, Theorem 3.2.3] every completely 0-simple semigroup is isomorphic to some $\mathcal{M}^0[G; I, \Lambda; P]$.

Semigroups of transformations

Let X be a non-empty set. The *symmetric group* S_X consists of all bijections from X to itself under composition of maps. The *full transformation semigroup* T_X consists of all maps from X into X under the operation of composition of maps. The *partial transformation semigroup* P_X consists of all partial maps of X, while the *symmetric inverse semigroup* I_X consists of all partial one-one maps of X. When $|X| = n$ we often write S_n, T_n, P_n and I_n in place of S_X, T_X, P_X and I_X respectively. When $|X| = n$ we often identify X with the set $X_n = \{1, 2, \ldots, n\}$. For every $\alpha \in T_X$ we define

$$\operatorname{im} \alpha = \{x\alpha : a \in X\}$$

and call this the *image* of the map α. Also we define

$$\ker \alpha = \{(x, y) \in X \times X : x\alpha = y\alpha\}$$

and call this the *kernel* of the map α. This is clearly an equivalence relation on the set X. We call the equivalence classes of $\ker \alpha$ the *kernel classes* of the map α. A *partition* of the set X is a family of pairwise disjoint, non-empty subsets of X whose union is X. Thus the kernel classes of a transformation in T_X are a partition of the set X. Often it will be convenient to write $\ker \alpha$ in terms of this partition rather than as a subset of $X \times X$. For example, the kernel of the element

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 2 & 3 \end{pmatrix} \in T_4$$

would be written as $\ker \alpha = \{\{1\}, \{2, 3\}, \{4\}\}$.

Given a partition $\cup_{i \in I} X_i$ of X we define the *weight* of the partition to be $|I|$. The *Stirling number of the second kind*, $S(n, r)$, is the number of partitions of $\{1,\ldots,n\}$ with r (non-empty) parts. This number is given by the following recursion formula:

$$S(n+1,k) = kS(n,k) + S(n,k-1)$$

where $S(1,0) = 0$ and $S(1,1) = 1$. Given a partition of the set X we call a system of distinct representatives of this family of sets a *transversal* of the family.

Direct and semidirect products

If S and T are semigroups then the set $S \times T$ with multiplication

$$(s_1,t_2)(s_2,t_2) = (s_1s_2, t_1t_2)$$

forms a semigroup that we call the *direct product* of S and T. More generally, let T and S be semigroups and ψ be a homomorphism of S into $\mathrm{End}(T)$. Denote by st the value of $\psi(s) \in \mathrm{End}(T)$ at $t \in T$. We view ψ as a left action of S on T, ${}^{s_1}({}^{s_2}t) = {}^{s_1s_2}t$, by endomorphisms ${}^s(t_1t_2) = {}^st_1{}^st_2$. The *semidirect product* of T and S over ψ is the semigroup $T \times_\psi S$ on the Cartesian product $T \times S$ with multiplication

$$(t_1,s_1)(t_2,s_2) = (t_1{}^{s_1}t_2, s_1s_2).$$

In the special case where $\psi(S) = \{1_{\mathrm{End}(T)}\} \leq \mathrm{End}(T)$ the semidirect product of T and S over ψ is equal to the direct product of S and T.

Free semigroups, monoids and presentations

Let A be a non-empty set. Let A^+ be the set of all finite, non-empty words in the alphabet A. With respect to the binary operation of juxtaposition of words the set A^+ forms a semigroup that we call the *free semigroup* on A. The set A is a generating set for A^+ and it is the unique minimal generating set of A^+. Adjoining an identity 1 to the free semigroup A^+ gives the *free monoid* which we denote A^*. We think of the identity of the free semigroup A^* as the empty word and sometimes denote it by ϵ. Every semigroup can be expressed as a quotient of a free semigroup by a congruence. If A is a finite alphabet and if we can find a finite set $R \subseteq A^+ \times A^+$ such that $S \cong A^+/\rho$, where ρ is the smallest congruence of A^+ containing R, then we say that S is finitely presented.

1.2 Graph theory preliminaries

We need some basic concepts from graph theory. The ideas presented here are standard and may be found in any introductory text on graph theory. See for example [51], [10], [8] and [15].

Subgraphs and isomorphisms

A *graph* Γ is a pair $(\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ is a set and $\mathcal{E}$ is a set of 2-subsets of $\mathcal{V}$. The set $\mathcal{V}$ is the set of *vertices* and the set $\mathcal{E}$ is the set of *edges*. Given a graph Γ we use $\mathcal{V}(\Gamma)$ and $\mathcal{E}(\Gamma)$ to denote the set of vertices and the set of edges, respectively, of the graph Γ. An edge $\{i, j\}$ is said to *join* the vertices i and j and this edge is denoted ij. The vertices i and j are called the *endvertices* of the edge ij. If $ij \in \mathcal{E}(\Gamma)$ we say that the vertices i and j are *adjacent* in the graph Γ. We say that i and j are *incident* to the edge ij. We say that two edges are adjacent if they have a common incident vertex.

We think of a graph as a collection of vertices, some of which are joined by edges, and as a result graphs are often represented as pictures. For example, the graph $\Gamma = (\mathcal{V}, \mathcal{E}) = (\{1,2,3,4\}, \{\{1,2\},\{2,3\},\{3,4\}\})$ is given in Figure 1.1.

The graph $\Gamma' = (\mathcal{V}', \mathcal{E}')$ is a *subgraph* of Γ if $\mathcal{V}' \subseteq \mathcal{V}$ and $\mathcal{E}' \subseteq \mathcal{E}$. If Γ' contains all the edges of Γ that join vertices of $\mathcal{V}'$ then we call Γ' the subgraph of Γ *induced* by $\mathcal{V}'$.

Given a subset $\mathcal{W}$ of $\mathcal{V}(\Gamma)$ we use $\Gamma - \mathcal{W}$ to denote the subgraph of Γ obtained by deleting the vertices $\mathcal{W}$ and all of the edges adjacent with them. Similarly, given a subset $\mathcal{F}$ of the edge set $\mathcal{E}(\Gamma)$ we use $\Gamma - \mathcal{F}$ to denote the subgraph obtained by deleting the edges $\mathcal{F}$. An *elementary contraction* of a graph Γ is obtained by identifying two adjacent vertices u and v, that is, by deleting u and v and replacing them by a single vertex w adjacent all to the vertices to which u or v were adjacent. A graph Γ is *contractible* to a graph Γ' if Γ' can be obtained from Γ by a finite sequence of elementary contractions. If the graph Γ is contractable to the graph Γ' then we call Γ' a *contraction* of Γ.

The graphs $\Gamma = (\mathcal{V}, \mathcal{E})$ and $\Gamma' = (\mathcal{V}', \mathcal{E}')$ are said to be *isomorphic* if there

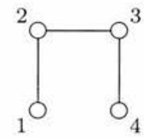


Figure 1.1: A graph.

is a bijection $\phi : \mathcal{V} \to \mathcal{V}'$ that preserves adjacency. That is, for all $x, y \in \mathcal{V}$, $xy \in \mathcal{E}$ if and only if $\phi(x)\phi(y) \in \mathcal{E}'$. We write $\Gamma \cong \Gamma'$ to mean that Γ and Γ' are isomorphic.

The *complete graph of order* n is defined to be the unique graph with n vertices and $\binom{n}{2}$ edges, so that there is an edge connecting every pair of vertices, and is denoted K_n.

The *degree* of a vertex v is the number of edges adjacent to it and will be denoted by $d(v)$. The set of vertices adjacent to a vertex v is called the *neighbourhood* of v and is denoted $N(v)$. More generally, if $\mathcal{W}$ is a subset of $\mathcal{V}$ then the neighbourhood of $\mathcal{W}$, denoted $N(\mathcal{W})$, is defined to be the set of all vertices of $\mathcal{V} \setminus \mathcal{W}$ which are neighbours of at least one vertex from $\mathcal{W}$. We call a graph k-*regular* if every vertex has degree k for some number k. A graph is *regular* if it is k-regular for some k.

Paths and connectedness

A *path* in a graph Γ is a set of vertices $\pi = \{v_0, v_1, \dots, v_n\}$ such that $v_{i-1}v_i$ belong to $\mathcal{E}(\Gamma)$ for all $1 \le i \le n$. We call v_0 and v_k the *initial* and *terminal* vertices, respectively, of the path π. A *walk* Γ is a sequence $(v_0, e_1, v_1, e_2, v_2, \dots, e_n, v_n)$ where e_i is the edge $v_{i-1}v_i$ for $i = 1, \dots, n$. A *trail* is a walk where all the edges are distinct. A trail whose endvertices coincide is called a *circuit*. A walk with at least three vertices, where all the vertices are distinct, and where the endvertices coincide, is called a *cycle*.

A graph Γ is called *connected* if for every pair $\{x, y\}$ of distinct vertices there is a path from x to y. The maximum connected subgraphs of a graph Γ are called the *connected components* of Γ. A *forest* is a graph with no cycles and a *tree* is a connected graph with no cycles. Therefore, the connected components of a forest are all trees. A *spanning tree* of a graph Γ is a subgraph $\mathcal{T}$ that is a tree and that contains every vertex of the graph Γ. Frequent use will be made of the following easy result.

Lemma 1.2. *Every connected graph contains a spanning tree.*

Of course, if Γ has n vertices and $\mathcal{T}$ is a spanning tree of Γ then $\mathcal{T}$ has n vertices and $n - 1$ edges.

A graph Γ is *bipartite* if $\mathcal{V}$ can be written as the disjoint union of two sets $\mathcal{V}_1$ and $\mathcal{V}_2$ in such a way that every edge in $\mathcal{E}$ has one vertex in $\mathcal{V}_1$ and the other in $\mathcal{V}_2$. We say that the bipartite graph $\Gamma = \mathcal{V}_1 \cup \mathcal{V}_2$ is *balanced* if $|\mathcal{V}_1| = |\mathcal{V}_2|$. Similarly, we say that the graph Γ is r-partite with vertex classes $\mathcal{V}_1, \dots, \mathcal{V}_r$ if

$V(\Gamma) = \mathcal{V}_1 \cup \ldots \cup \mathcal{V}_r$, and $\mathcal{V}_i \cap \mathcal{V}_j = \varnothing$ whenever $i \neq j$ and no edge joins two vertices of the same class. A *complete r-partite graph* is denoted $K_{n_1,\ldots,n_r}$, it has n_i vertices in the ith class and contains all edges joining vertices in distinct classes.

In the above definition of graph we do not allow multiple edges or loops (an edge joining a vertex to itself). We call a graph that is allowed multiple edges and loops a *multigraph.* If the edges are ordered pairs (rather than two-sets) then we get the notion of a *digraph* (directed graph) and *directed multigraph.* The notions above for graphs, such as paths and walks, carry over to the context of multigraphs and digraphs in a natural way.

Hamiltonian graphs

A cycle containing all the vertices of a graph is said to be a *Hamiltonian cycle.* A *Hamiltonian path* is a path containing all the vertices of a graph. A graph containing a Hamiltonian cycle is said to be *Hamiltonian.* No efficient algorithm is known for constructing a Hamiltonian cycle, though neither is it known that no such algorithm exists. On the other hand, some sufficient conditions for a graph to be Hamiltonian are known. The following result gives a sufficient condition for a bipartite graph to be Hamiltonian.

Theorem 1.3 (Moon and Moser, [75]). *If $\Gamma = X \cup Y$ is a bipartite graph with $|X| = |Y| = n$ such that any non-adjacent pair of vertices $(x, y) \in X \times Y$ satisfies $d(x) + d(y) \geq n + 1$, then Γ is Hamiltonian.*

Matchings and Hall's theorem

A subset $\mathcal{F}$ of $\mathcal{E}(\Gamma)$ is called *independent* if no two edges have a vertex in common. Similarly, a subset $\mathcal{V}'$ of $\mathcal{V}(\Gamma)$ is called independent of no two vertices in $\mathcal{V}'$ are adjacent. A *matching* in a graph is a set of independent edges. A *perfect matching* is a matching on $|\mathcal{V}|/2$ edges. In particular, in a bipartite graph $\Gamma = A \cup B$ associated with any perfect matching is a bijection $\pi : A \to B$ that satisfies $\{x, x\pi\} \in \mathcal{E}(\Gamma)$ for all $x \in A$. Note that if $\Gamma = A \cup B$ has a perfect matching then $|A| = |B|$.

Let $A_1, \ldots, A_n$ be sets. A *system of distinct representatives* (SDR) for these sets is an n-tuple $(x_1, \ldots, x_n)$ of element with the properties:

(i) $x_i \in A_i$ for $i = 1, \ldots, n$;

(ii) $x_i \neq x_j$ for $i \neq j$.

Hall's marriage theorem gives necessary and sufficient conditions for a family $(A_1, \ldots, A_n)$ of finite sets to have a SDR.

Theorem 1.4 (Hall's theorem, [50]). *The family* $(A_1, \ldots, A_n)$ *of finite sets to have a SDR if and only if*

$$\left| \bigcup_{j \in J} A_j \right| \geq |J| \text{ for every } J \subseteq \{1, \ldots, n\}. \tag{1.1}$$

When a family of sets $(A_1, \ldots, A_n)$ satisfies condition (1.1) we say that it satisfies *Hall's condition.* A family $\mathcal{A} = (A_1, \ldots, A_n)$, where $A_i \subseteq X$ for all i, is naturally identifiable with a bipartite graph with vertex classes $\mathcal{V}_1 = \mathcal{A}$ and $\mathcal{V}_2 = X$ where $A_i \in \mathcal{A}$ is joined to $x \in X$ if and only if $x \in A_i$. A system of distinct representatives is then just a perfect matching in this bipartite graph. In this context Hall's marriage theorem becomes.

Theorem 1.5. *The bipartite graph* $G = X \cup Y$ *has a perfect matching if and only if* $|N(A)| \geq |A|$ *for all subsets* A *of* X.

A *colouring* of a graph is an assignment of colours to the vertices such that adjacent vertices have distinct colours. A k-colouring of Γ is a function $c : \mathcal{V}(\Gamma) \to \{1, 2, \ldots, k\}$ such that for each j the set $c^{-1}(j)$ is independent. The *chromatic number* $\chi(\Gamma)$ of the graph Γ is the minimal number of colours in a vertex colouring of the graph Γ.

The *dual* of the graph Γ is the graph $D(\Gamma)$ with vertex set $\mathcal{V}(\Gamma)$ and $ij \in \mathcal{E}(D(\Gamma))$ if and only if $ij \notin \mathcal{E}(\Gamma)$.

Part I

Finite semigroup theory

Chapter 2

Generating sets for completely 0-simple semigroups using bipartite graphs

2.1 Finite semigroups and their generating sets

It is often convenient to give a finite semigroup S in terms of a set of generators A. In many cases this set may be chosen to have considerably fewer elements than S itself. For example, the full transformation semigroup T_n has n^n elements while it may be generated by just three transformations. In particular, the transposition $(1\ 2)$, the n-cycle $(1\ 2\ \dots\ n)$, and any transformation α satisfying $|\operatorname{im}\alpha| = n-1$, together will generate T_n (see [57, Exercise 1.7]).

In this chapter we will be concerned with the problem of finding "small" generating sets for finite semigroups. Given a semigroup S we will use $\operatorname{rank}(S)$ to denote the minimum cardinality of a generating set for S. In other words:

$$\operatorname{rank}(S) = \min\{|A| : \langle A\rangle = S\}.$$

Our interest is in determining $\operatorname{rank}(S)$ and, whenever possible, in describing all generating sets with this size. We will call any generating set of S with size $\operatorname{rank}(S)$ a *basis* of the semigroup S.

The ranks of a wide number of finite groups are well known. In a finite group G every subsemigroup is a subgroup (since for any $g \in G$ we have $g^m = g^{-1}$ for some positive number m) and thus, given a subset A of G, the subsemigroup of G generated by A is equal to the subgroup of G generated by A. There is, therefore, no distinction between the *group rank* and the *semigroup rank* of G. For infinite groups this is not necessarily the case. For example, the infinite cyclic group $\mathbb{Z}$ has rank 2 as a semigroup but rank 1 as a group. It is well known that the symmetric group satisfies $\operatorname{rank}(S_n) = 2$ for $n \geq 3$, as does the alternating group A_n for $n \geq 4$. In fact, any finite non-abelian simple group G has rank 2. This result is a consequence of the classification of finite simple groups. The rank of any non-trivial finite general linear group is also known to equal 2 (see for example [94]).

The function $\operatorname{rank} : \mathcal{S} \to \mathbb{N}$, from the class (pseudo-variety) of all finite semigroups to the natural numbers, does not behave well with respect to taking subsemigroups. For example, by Cayley's theorem, every finite semigroup is embeddable in some finite full transformation semigroup T_n while, as already mentioned, $\operatorname{rank}(T_n) = 3$ for all $n \geq 3$. On the other hand, if T is the image of S under a homomorphism then it is clear that $\operatorname{rank}(S) \geq \operatorname{rank}(T)$. This is because of the following simple observation.

Lemma 2.1. *Let S and T be semigroups, let A be a subset of S and let $\phi : S \to T$ be an epimorphism. If A generates S then $A\phi$ generates T. In particular*

$|A\phi| \leq |A|$ *and* $\text{rank}(S) \geq \text{rank}(T)$.

In [23] a group G is defined to be *generator critical* if all of its proper homomorphic images H satisfy $\text{rank}(G) > \text{rank}(H)$. The idea being that if a group is *not* generator critical then we may factor down to an "easier" group with the same rank. Exactly the same idea carries over to semigroups and the general idea of studying large homomorphic images of S in order to determine its rank will be a re-occurring theme throughout this chapter.

In terms of semigroup theory, the question of rank has been considered mainly for various semigroups of transformations. The theory of transformation semigroups is one of the oldest and most developed within semigroup theory. In [58] Howie argues that

> "It is this connection with maps (arising from the associative axiom) that is the strongest reason why semigroups are more important both theoretically and in applications that the various non-associative generalizations of groups".

Early work on generators and relations in transformation semigroups was carried out by Aĭzenštat in [2] and [3]. In [37] Gomes and Howie prove that the semigroup Sing_n of of all singular self-maps of X_n satisfies $\text{rank}(\text{Sing}_n) = n(n-1)/2$. In the same paper they also consider the semigroup $SP_n \leq I_n$, of all proper subpermutations of X_n, proving that $\text{rank}(SP_n) = n+1$. In [59] Howie and McFadden generalized the above result for the semigroup of singular mappings by considering a general two-sided ideal of T_n. These ideals have the form:

$$K(n,r) = \{\alpha \in T_n : |\operatorname{im} \alpha| \leq r\}$$

where $1 \leq r \leq n$ (we will see the reason for this in Proposition 2.14). In particular, in [59] it is shown that $\text{rank}(K(n,r)) = S(n,r)$: the Stirling number of the second kind. Garba, in [34], considered the semigroup of all partial transformations P_n on the set X_n and showed $\text{rank}(KP(n,r)) = S(n+1, r+1)$ where

$$KP(n,r) = \{\alpha \in P_n : |\operatorname{im} \alpha| \leq r\}.$$

In [36] he also generalised Gomes and Howie's result for SP_n by showing $\text{rank}(L(n,r)) = \binom{n}{r} + 1$, where

$$L(n,r) = \{\alpha \in I_n : |\operatorname{im} \alpha| \leq r\}.$$

Various order preserving versions of the examples above have also been considered. Originally in [4] Aĭzenštat considered the semigroup of order preserving transformations:

$$O_n = \{\alpha \in \mathrm{Sing}_n : (\forall x, y \in X_n)\ x \leq y \Rightarrow x\alpha \leq y\alpha\}.$$

She showed that it is idempotent generated and that it has a uniquely determined irreducible set of idempotent generators (namely the identity along with all the idempotents e that satisfy $|\operatorname{im} e| = n-1$). This result was later reproven by Howie in [38]. We will see more on idempotent generating sets in subsequent chapters. Also, in [38] it was shown that the semigroup O_n has rank n. In the same paper the semigroup of partial order preserving transformations of X_n (excluding the identity map):

$$PO_n = O_n \cup \{\alpha : \mathrm{dom}(\alpha) \subsetneq X_n, (\forall x, y \in \mathrm{dom}(\alpha))\ x \leq y \Rightarrow x\alpha \leq y\alpha\}$$

was shown to have rank $2n-1$ and the strictly partial order preserving transformations:

$$SPO_n = PO_n \setminus O_n$$

were shown to have rank $2n-2$. Also, in a series of papers [68], [67] and [66] Levi and Seif have considered semigroups generated by transformations of prescribed partition type. These semigroups are closely related to the S_n-normal semigroups introduced in [65].

Given an arbitrary finite semigroup S, if A generates S and J_M is some maximal $\mathcal{J}$-class of S then $A \cap J_M$ must generate the principal factor $J_M{}^*$. As a consequence, the rank of S is equal to at least the sum of the ranks of the principal factors that correspond to the maximal $\mathcal{J}$-classes of S. If S happens to be generated by the elements of its maximal $\mathcal{J}$-classes then $\mathrm{rank}(S)$ is precisely equal to this sum. In fact, this is a property that is shared by the majority of the semigroups described above. As a consequence, in each case the rank of the semigroup in question is equal to the rank of a corresponding completely 0-simple semigroup. In this way, these results act as motivation for finding a general formula for the rank of an arbitrary finite completely 0-simple semigroup.

The first occurrence of a formula for the rank of a completely 0-simple semigroup can be found in [37] where, in order to find the rank of the semigroup SP_n, the authors give an expression for the rank of an arbitrary Brandt semigroup $B(G, \{1, \ldots, n\})$ in terms of its dimension n, and of the rank of the underlying group G. In another paper [81] the author considers a class of completely 0-simple semigroups he calls connected, a restriction on the form of the matrix P which in particular is satisfied by all completely simple semigroups, and gives a formula for the rank of an arbitrary connected completely 0-simple semigroup.

In this chapter we will build on the ideas of [81] giving a general formula for the rank of an arbitrary completely 0-simple semigroup in terms of the group G, the size of the index sets I and Λ, the number of "components" in the matrix P, and a special term $r_{\min}$ that will be defined. In §2.2 some preliminary results are introduced then in §2.3 the special case of combinatorial completely 0-simple semigroups (those whose maximal subgroups are trivial) is considered. Graph theoretic methods for working with completely 0-simple semigroups are introduced in §2.4. Results for connected completely 0-simple semigroups are given in §2.5 and in §2.6–2.9 the general case is considered and the main results of the chapter are presented. A normalization theorem is the subject of §2.10 and finally, in §2.11–2.13, several applications of the main results are discussed.

2.2 Preliminaries

Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a finite completely 0-simple semigroup. We will use $\mathcal{R}_i$, $\mathcal{L}_\lambda$ and $\mathcal{H}_{i\lambda}$ to denote the $\mathcal{R}$, $\mathcal{L}$ and $\mathcal{H}$-classes indexed by $i \in I$, $\lambda \in \Lambda$ and $(i, \lambda) \in I \times \Lambda$ respectively. We can only hope to answer our question "modulo groups" and the concept of *relative rank* gives us a way of accomplishing this. Given a subset A of a semigroup S, we define the *relative rank* of S modulo A as the minimum number of elements of S that need to be added to A in order to generate the whole of S:

$$\operatorname{rank}(S : A) = \min\{|X| : \langle A \cup X \rangle = S\}.$$

Example 2.2. From the discussion at the beginning of Section 2.1 we conclude that $\operatorname{rank}(T_n : S_n) = 1$.

Since all the semigroups we consider here have a zero we will always include the zero in any given subsemigroup. As a consequence of this by $\langle X \rangle$ we will mean all the elements that can be written as products of elements of X, plus zero if necessary. This is really just a matter of convenience and we do not lose anything by doing it. Without this convention we would have to deal with the cases $0 \in (S \setminus \{0\})^2$ and $0 \notin (S \setminus \{0\})^2$ separately with the rank differing by 1 each time.

The following lemma states the obvious fact that a generating set for $S = \mathcal{M}^0[G; I, \Lambda; P]$ must intersect every $\mathcal{R}$- and every $\mathcal{L}$-class of S.

Lemma 2.3. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a completely* 0*-simple semigroup. Then* $\operatorname{rank}(S) \geq \max(|I|, |\Lambda|)$.

As a direct consequence of Green's Lemmas (see Chapter 1) it follows that if A generates every element of a single group $\mathcal{H}$-class and at least one element in every other $\mathcal{H}$-class then this is enough to say that A generates the whole of S. More precisely:

Lemma 2.4. *[81, Lemma 3.7] Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a completely* 0*-simple semigroup, let $H_{i\lambda}$ be a group and let $A \subseteq S$. If $H_{i\lambda} \subseteq \langle A\rangle$ and $\langle A\rangle \cap H_{j\mu} \neq \varnothing$ for all $j \in I, \mu \in \Lambda$ then $S = \langle A\rangle$.*

The result above may be thought of as analogous to the following situation in group theory. Let G be a group and let N be a normal subgroup of G. Let A be a subset of G. If $\langle A\rangle \supseteq N$ and $\langle A/N\rangle = G/N$ then $\langle A\rangle = G$. Here N is playing the role of $H_{i\lambda}$ and the cosets of N in G play the role of the non-trivial $\mathcal{H}$-classes of S.

Roughly speaking, if S is a completely 0-simple semigroup and A is a generating set for S then every generator $a \in A$ makes a two-fold contribution. Firstly, the generator contributes to generating at least one element in every $\mathcal{H}$-class of S. Secondly, each generator contributes to generating the underlying group.

2.3 Rectangular 0-bands

A *rectangular* 0*-band*, denoted by $S = \mathcal{M}^0[\{1\}; I, \Lambda; P]$, is a 0-Rees matrix semigroup over the trivial group. Understanding the generating sets of rectangular 0-bands will give us a useful first step towards understanding generating sets of Rees matrix semigroups over non trivial groups. Since the middle component of every triple equals 1 we can effectively ignore it and consider the semigroup of pairs $S = (I \times \Lambda) \cup \{0\}$ with $I = \{1, 2, \ldots, m\}$ and $\Lambda = \{1, 2, \ldots, n\}$, P a regular $n \times m$ matrix over $\{0, 1\}$, whose multiplication is given by

$$(i, \lambda)(j, \mu) = \begin{cases} (i, \mu) & \text{if } p_{\lambda j} = 1 \\ 0 & \text{if } p_{\lambda j} = 0 \end{cases}$$

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

Associated with any completely 0-simple semigroup $S = \mathcal{M}^0[G; I, \Lambda; P]$ is the rectangular 0-band given by replacing all the non-zero entries in the matrix P by the symbol 1 and by replacing G by the trivial group.

Definition 2.5. Given $S = \mathcal{M}^0[G; I, \Lambda; P]$ by the *natural rectangular 0-band homomorphic image of S* we mean $S/\mathcal{H}$. This semigroup can be concretely represented as $T = \mathcal{M}^0[\{1\}; I, \Lambda; Q]$ where $q_{\lambda i} = 1$ if $p_{\lambda i} \neq 0$, and $q_{\lambda i} = 0$ otherwise. We will use $\natural$ to denote the corresponding epimorphism from S to T with $0\natural = 0$ and $(i, g, \lambda)\natural = (i, \lambda)$.

We call two $n \times m$ matrices A and B over $\{0, 1\}$ *equivalent*, writing $A \sim B$, if B can be obtained from A by permuting its rows and columns. Clearly $\sim$ is an equivalence relation on the set of $n \times m$ matrices over $\{0, 1\}$ and, as a special case of Theorem 2.54 below, two matrices are equivalent if and only if the rectangular 0-bands that they correspond to are isomorphic.

Definition 2.6. Let P be an $n \times m$ matrix over $\{0, 1\}$. Then we use $P[i_1, \ldots, i_l][j_1, \ldots, j_k]$ to denote the submatrix of P obtained by deleting all elements with first coordinate in the set $\{i_1, \ldots, i_l\}$ or second coordinate in the set $\{j_1, \ldots, j_k\}$. We use 0 when no rows or columns are to be deleted. For example $P[0][1]$ means leave the rows alone but delete column 1.

First we will show that given a non-square matrix we can always delete a row or a column while maintaining regularity.

Lemma 2.7. *Let P be a regular $n \times m$ matrix over $\{0, 1\}$.*

(i) If $m > n$ then there exists $j \in \{1, \ldots, m\}$ such that $Q = P[0][j]$ is regular.

(ii) If $m < n$ then there exists $i \in \{1, \ldots, n\}$ such that $Q = P[i][0]$ is regular.

(iii) If $n = m$ then there exists $j \in \{1, \ldots, n\}$ such that $P[0][j]$ is regular if and only if there exists $i \in \{1, \ldots, n\}$ such that $P[i][0]$ is regular which is the case if and only if $P \not\sim I_n$ (the $n \times n$ identity matrix).

Proof. (i) Suppose otherwise, so that for all $j \in \{1, \ldots, m\}$ the matrix $P[0][j]$ is not regular. Then for each $j \in \{1, \ldots, m\}$ there is a row that has 1 in the jth position and zeros everywhere else. All of these rows are distinct and there are m of them which contradicts the fact that $m > n$. (ii) Use a symmetric argument to that of part (i). (iii) By the same argument as in part (i) it follows that if $P[0][i]$ is not regular for all $i \in \{1, \ldots, n\}$ then $P \sim I_n$. Conversely, if $P \sim I_n$ it is clear that for any $i \in \{1, \ldots, n\}$, $P[0][i]$ is not regular. □

Corollary 2.8. *Let P be a regular $n \times n$ matrix over $\{0, 1\}$ such that $P \not\sim I_n$. Then there exist $i, j \in \{1, \ldots, n\}$ such that $P[0][j], P[i][0]$ and $P[i][j]$ are all regular.*

Proof. By Lemma 2.7 we can find $i, j \in \{1, \ldots, n\}$ so that $P[0][j]$ and $P[i][0]$ are regular. Hence $P[i][j]$, which is their intersection, is clearly regular. □

The *Brandt* semigroup $B = B(G, \{1, \ldots, n\})$ is the Rees matrix semigroup $\mathcal{M}^0[G; I, I; P]$ where $P \sim I_n$, the $n \times n$ identity matrix, and $I = \{1, \ldots, n\}$. When G is the trivial group $B(G, I)$ is a rectangular 0-band that we call the *aperiodic Brandt semigroup* and denote by B_n.

Lemma 2.9. *The aperiodic Brandt semigroup B_n has rank n.*

Proof. ($\geq$) By Lemma 2.3 $\text{rank}(B_n) \geq n$. ($\leq$) The set $\{(1,2), (2,3), \ldots, (n-1, n), (n, 1)\}$ generates B_n. Indeed, if $(x, y) \in B_n$ then:

$$(x,y) = \begin{cases} (x, x+1)(x+1, x+2) \ldots (y-1, y) & \text{if } y > x \\ (x, x+1)(x+1, x+2) \ldots (n-1, n)(n, 1)(1, 2) \ldots (y-1, y) & \text{if } y \leq x. \end{cases}$$

□

In fact, it is fairly easy to describe all the bases of the aperiodic Brandt semigroup B_n and we will do so at the end of §2.4. Corollary 2.8 forms the basis of the inductive step that proves the following result.

Theorem 2.10. *Let $S = \mathcal{M}^0[\{1\}; I, \Lambda; P]$ be an $m \times n$ rectangular 0-band. Then*

$$\text{rank}(S) = \max(m, n).$$

Proof. It follows from Lemma 2.3 that $\text{rank}(S) \geq \max(m, n)$. Now we have to show that we can always find a generating set of this size. We must prove:

> $\text{U}(m, n)$: If S is an $m \times n$ rectangular 0-band then S has a generating set with cardinality $\max(m, n)$.

First we consider the case where $m = n$ and use induction on n. $\text{U}(1, 1)$ holds trivially. Suppose $\text{U}(k, k)$ holds and let T be a $(k+1) \times (k+1)$ rectangular 0-band with $I = \Lambda = \{1, 2, \ldots, k+1\}$, and underlying matrix P. If $P \sim I_{k+1}$ then $\text{U}(k+1, k+1)$ holds by Lemma 2.9. If $P \not\sim I_{k+1}$ then by Corollary 2.8 we can suppose without loss of generality that the submatrices $M = P[k+1][0]$, $N = P[0][k+1]$ and $O = P[k+1][k+1]$ are all regular. Let T_M, T_N and T_O be the sub-rectangular 0-bands corresponding to these regular matrices. By the inductive hypothesis we can find $A \subseteq T_O$ such that $|A| = k$ and $\langle A \rangle = T_O$. Now let $B = A \cup \{(k+1, k+1)\} \subseteq T$. Clearly $|B| = k+1$ and we also claim that $\langle B \rangle = T$. Indeed, we have

$$\langle B \rangle = \langle A \cup \{(k+1,k+1)\} \rangle = \langle \langle A \rangle \cup \{(k+1,k+1)\} \rangle = \langle T_0 \cup \{(k+1,k+1)\} \rangle.$$

We are left to show $\{(k+1, i), (i, k+1) : i \in \{1, \dots, k\}\} \subseteq \langle B \rangle$. Let $j \in \{1, \dots, k\}$. By the regularity of M and N we can find $v, l \in \{1, \dots, k\}$ such that $p_{(k+1),l} = p_{v,(k+1)} = 1$. Then we have

$$(k+1, j) = (k+1, k+1)(l, j), \quad (j, k+1) = (j, v)(k+1, k+1)$$

with $(l, j), (j, v) \in T_0$ which completes the inductive step.

Now we consider the case where T is an $m \times n$ rectangular 0-band with, say, $n > m$. By repeated application of Lemma 2.7 without loss of generality we can suppose that $Q = P[m+1, m+2, \dots, n][0]$ is regular. Let T_Q be the sub-rectangular 0-band corresponding to Q. Since Q is an $m \times m$ matrix by the previous case we can find $A \subseteq T_Q$ with $|A| = m$ and $\langle A \rangle = T_Q$. Now let $R = \{(1, \lambda) : \lambda = m+1, \dots, n\}$ and $B = A \cup R \subseteq T$. Clearly $|B| = n$ and we also claim that $\langle B \rangle = T$. Indeed, we have

$$\langle B \rangle = \langle A \cup R \rangle = \langle \langle A \rangle \cup R \rangle = \langle T_Q \cup R \rangle.$$

We are left to show $\{(i, j) : i \in \{1, \dots, m\}, j \in \{m+1, \dots, n\}\} \subseteq \langle B \rangle$. By the regularity of Q we can find $x \in \{1, \dots, m\}$ such that $p_{x1} = 1$ and we conclude that $(i, j) = (i, x)(1, j)$ where $(i, x) \in T_Q$ and $(1, j) \in R$. □

Using just this simple result we may now determine the rank of a special class of completely 0-simple semigroup.

Definition 2.11. A semigroup S is called *idempotent generated* if $\langle E(S) \rangle = S$.

Lemma 2.12. *Let S be a finite idempotent generated semigroup and let A be a subset of S. If A has non-trivial intersection with every $\mathcal{H}$-class of S then A generates S.*

Proof. Since $S = \langle E(S) \rangle$ it is sufficient to prove $E(S) \subseteq \langle A \rangle$. Let $e \in E(S)$ and let $b \in A \cap H_e$. Since H_e is a finite group with identity e it follows that $b^i = e$ for some $i \in \mathbb{N}$, the smallest such i just being the order of b in the group H_e. It follows that $e = b^i \in \langle A \rangle$ and since e was arbitrary that $E(S) \subseteq \langle A \rangle$. □

Combining this with Theorem 2.10 gives:

Theorem 2.13. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a finite completely 0-simple semigroup. If S is idempotent generated then* $\operatorname{rank}(S) = \max(|I|, |\Lambda|)$.

Proof. The fact that $\operatorname{rank}(S) \geq \max(|I|, |\Lambda|)$ follows from Lemma 2.3. For the converse let $T = S\natural$, the natural rectangular 0-band homomorphic image of S. By Theorem 2.10 we can find a generating set A for T with size $\max(|I|, |\Lambda|)$. The pre-image of A under the map $\natural$ is a union of $\mathcal{H}$-classes of S. Let B be a transversal of this set of $\mathcal{H}$-classes. Since $\langle A \rangle = T$ it follows that $\langle B \rangle$ has non-trivial intersection with every (non-zero) $\mathcal{H}$-class of S and so B generates S, by Lemma 2.12, and $|B| = |A| = \max(|I|, |\Lambda|)$. □

This result is not quite as obscure as it might at first seem. Many naturally occurring semigroups are idempotent generated and when, in addition to this, they are generated by the elements in their maximal $\mathcal{J}$-classes, determining their rank just reduces to the problem of counting the number of $\mathcal{R}$- and $\mathcal{L}$-classes. The full transformation semigroup T_n provides us with a family of examples of this kind.

Proposition 2.14. *Let $n \in \mathbb{N}$ and let $1 \leq r < n$. Let $\alpha, \beta \in T_n$.*

(i) Green's relations are given by:

(a) $\alpha \mathcal{R} \beta \Leftrightarrow \operatorname{im} \alpha = \operatorname{im} \beta$;

(b) $\alpha \mathcal{L} \beta \Leftrightarrow \ker \alpha = \ker \beta$;

(c) $\alpha \mathcal{D} \beta \Leftrightarrow |\operatorname{im} \alpha| = |\operatorname{im} \beta|$.

(ii) $K(n, r) = \langle E(D_r) \rangle$.

(iii) The number of $\mathcal{R}$-classes in $D_r = \{\alpha \in T_n : |\operatorname{im} \alpha| = r\}$ is the Stirling number of the second kind $S(n, r)$.

(iv) The number of $\mathcal{L}$-classes in D_r is $\binom{n}{r}$.

(v) The $\mathcal{H}$-class indexed by the image I and the kernel K is a group if and only if I is a transversal of K.

Proof. (i) See [57, Exercise 1.16]. (ii) By [57, Theorem 6.3.1] the semigroup Sing_n is regular and idempotent generated. By [57, Exercise 6.12] if S is a semiband then every element $a \in S$ is expressible as a product of idempotents in J_a. Also, from [57, Lemma 6.3.2] it follows that $\langle D_r \rangle = K(n, r)$. The result is an immediate consequence of these three facts. (iii) and (iv) are immediate consequences of (i). (v) See [57, Exercise 1.18]. □

These facts, along with Theorem 2.13 allow us to determine the rank of the semigroup $K(n,r)$. This result was originally proven in [59, Theorem 5], where they also determined the so called idempotent rank of the semigroup. We will see more about this in the following chapter.

Theorem 2.15. *Let $n \in \mathbb{N}$ and let $1 < r < n$. Then:*

$$\operatorname{rank}(K(n,r)) = S(n,r).$$

Proof. It follows from Proposition 2.14 that $\operatorname{rank}(S) = \operatorname{rank}(K(n,r)/K(n,r-1))$ and that $K(n,r)/K(n,r-1)$ is idempotent generated. Applying Theorem 2.13 to the idempotent generated completely 0-simple semigroup $K(n,r)/K(n,r-1)$ gives:

$$\operatorname{rank}(K(n,r)) = \operatorname{rank}(K(n,r)/K(n,r-1)) = \max(S(n,r), \binom{n}{r}) = S(n,r).$$

□

Theorem 2.13 may be applied to a number of other examples. In particular the exact analogue of the above result may be proven for the ideals of $\operatorname{End}(V)$ where V is a finite vector space. See Section 3.5.2 for more details on this.

Another consequence of Theorem 2.10 is that with $S = \mathcal{M}^0[G; I, \Lambda; P]$ we have

$$\operatorname{rank}(S) \leq \operatorname{rank}(G) + \max(|I|, |\Lambda|).$$

We obtain this bound by joining together a generating set for $T = S\natural$ and a generating set for G. In each case, exactly where the answer lies between $\max(|I|, |\Lambda|)$ and $\operatorname{rank}(G) + \max(|I|, |\Lambda|)$ will depend on the "contribution" that is made by the idempotents of S.

2.4 Finite 0-simple semigroups and their associated graphs

Given an element $(i, g, \lambda) \in S = \mathcal{M}^0[G; I, \Lambda; P]$ we may visualise this triple as two vertices i and λ joined by a directed edge labelled with g:

$$(i, g, \lambda) \equiv i \xrightarrow{g} \lambda .$$

Taking this idea further, we may wish to view composition of elements of S as composition of such paths so that:

$$(i,g,\lambda)\circ(j,h,\mu) \equiv\ i \xrightarrow{\ g\ } \lambda\ \circ\ j \xrightarrow{\ h\ } \mu$$

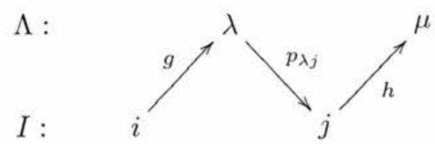
and since (provided $p_{\lambda j} \neq 0$) we have:

$$(i,g,\lambda)(j,h,\mu) = (i, gp_{\lambda j}h, \mu)$$

we amend our diagram to give:

$$i \xrightarrow{\ g\ } \lambda \xrightarrow{p_{\lambda j}} j \xrightarrow{\ h\ } \mu\ .$$

Grouping together elements of I and those of Λ gives:



which is starting to take the form of a directed bipartite graph with edges labelled by elements of G.

Such representations of completely 0-simple semigroups have been exploited with success in the past. The first place that such an idea appears in the literature is in [41]. In this paper Graham uses the graph theoretic approach to describe all maximal nilpotent subsemigroups of $\mathcal{M}^0[G; I, \Lambda; P]$ (a semigroup T is nilpotent if for some $n \in \mathbb{N}$ we have $T^n = \{0\}$). Moreover, necessary and sufficient conditions are given for a completely 0-simple semigroup to have a unique maximal nilpotent subsemigroup. Secondly, in the same paper a new normal form is introduced for completely 0-simple semigroups (see Section 2.10 for more details on this). This normal form is used to give a general description of the form that the maximal subsemigroups of an arbitrary finite completely 0-simple semigroup must take. This "local" result was later successfully used by Graham, Graham and Rhodes in [40] to give the form of the maximal subsemigroups of arbitrary finite semigroups. In [56] the bipartite graph representation was used to describe the subsemigroup generated by the idempotents of a completely 0-simple semigroup S and in [54] Houghton considered the homological properties of these graphs.

Here we will define three graphs, each with a different purpose. The first helps us find what a given subset of S generates. The second facilitates the study of $\langle E(S)\rangle$ and the third gives us the concept of connectedness in S.

Graph 1: $\boldsymbol{\Delta(S:A)}$

Given $S = \mathcal{M}^0[G; I, \Lambda; P]$ and $A \subseteq S$ where $0 \notin A$ we define a bipartite digraph $\Delta(S : A)$ with labelled edges in the following way. The vertex set of $\Delta(S : A)$ is $I \cup \Lambda$, where I and Λ are assumed to be disjoint. Edges from I to Λ represent elements of A and edges from Λ to I represent idempotents of S in the following way:

(i) corresponding to each $a = (i, g, \lambda) \in A$ there is an edge $i \xrightarrow{g} \lambda$ labelled with g;

(ii) corresponding to each non-zero entry $p_{\mu j} \in P$ there is an edge $\mu \xrightarrow{p_{\mu j}} j$ labelled with $p_{\mu j}$.

Note that the graph $\Delta(S : A)$ is allowed to have multiple edges, so when describing a path in this graph it is not enough to just give an ordered list of vertices that the path is to traverse.

Definition 2.16. Let $f = i \xrightarrow{g} \lambda$ be an edge from I to Λ and $e = \mu \xrightarrow{p_{\mu j}} j$ be an edge from Λ to I in $\Delta(S : A)$. We define the functions V and W:

$$V(f) = g, \quad V(e) = p_{\mu j}, \quad W(f) = (i, g, \lambda) \in A.$$

Let $x, y \in I \cup \Lambda$ and let $p = (e_1, e_2, \ldots, e_k)$ be a directed path in $\Delta(S : A)$ starting in x and ending in y. Then we write

$$V(p) = V(e_1)V(e_2)\ldots V(e_k) \in G,$$

and call this the *value* of the path p. We will use $\mathcal{P}_{x,y}$ with $x, y \in I \cup \Lambda$ to denote the set of all paths starting at x and ending at y in $\Delta(S : A)$ and define $V_{x,y} = \{V(p) : p \in P_{x,y}\}$. We call paths that start in I and end in Λ the *valid paths*. For every valid path $p = (f_1, e_1, f_2, e_2, \ldots, f_{k-1}, e_{k-1}, f_k)$ we write

$$W(p) = W(f_1)W(f_2)W(f_3)\ldots W(f_k) \in S.$$

Clearly if p is a valid path from i to λ then

$$W(p) = (i, V(p), \lambda).$$

There is a clear correspondence between non-zero products of elements of A and valid paths in the graph $\Delta(S:A)$. As a consequence of this correspondence we have the following straightforward lemma:

Lemma 2.17. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ and let $A \subseteq S$. If R is the set of valid paths in $\Delta(S:A)$ then*

$$\langle A \rangle = \{W(p) : p \in R\} \cup \{0\}.$$

Proof. This is obvious from the definitions. □

Graph 2: $\Delta(P)$

When $A = E(S)$ the graph $\Delta(S:A) = \Delta(S:E(S))$ takes a particularly nice form. Since $E(S) = \{(i, {p_{\lambda i}}^{-1}, \lambda) : p_{\lambda i} \neq 0\} \cup \{0\}$ every edge $f = i \xrightarrow{p_{\lambda i}} \lambda$ has a corresponding reverse edge $e(f) = \lambda \xrightarrow{{p_{\lambda i}}^{-1}} i$. In this situation we can simplify the graph $\Delta(S:E(S))$ in the following way. Let $\Delta(P)$ denote the underlying undirected graph of $\Delta(S:E(S))$ noting that $\Delta(P)$ has precisely one edge corresponding to each non-zero $p_{\lambda i}$ of P. In $\Delta(P)$ the edges are unlabelled but we will still assign values to the paths through the graph. The value of the path $\pi = z_1 \to z_2 \to \ldots \to z_t$ is defined to be

$$V(\pi) = \phi(z_1, z_2)\phi(z_2, z_3)\ldots\phi(z_{t-1}, z_t)$$

where

$$\phi(i, \lambda) = p_{\lambda i}^{-1}, \quad \phi(\lambda, i) = p_{\lambda i}, \quad i \in I, \quad \lambda \in \Lambda,$$

and $\mathcal{P}_{x,y}$ and $V_{x,y}$ have the same meaning as before. Note that the graph $\Delta(P)$ does not have multiple edges and so paths in the graph are uniquely determined by ordered lists of vertices. Given two vertices x and y in $\Delta(P)$ we write $x \bowtie y$ if there is a path from x to y in the graph $\Delta(P)$. This connectedness relation $\bowtie$ on the graph $\Delta(P)$ is an equivalence relation on the set $I \cup \Lambda$ and in [56] Howie proves the following result:

Theorem 2.18. *[56, Theorem 1] Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a completely 0-simple semigroup, let E be the set of idempotents in S. Then*

$$\langle E \rangle = \{(i, a, \lambda) \in S : i \bowtie \lambda \text{ and } a \in V_{i,\lambda}\} \cup \{0\}.$$

Graph 3: $\Gamma(\mathbb{H}_S)$

Given $C \subseteq I \times \Lambda$ we will let $\Gamma(C)$ denote the undirected graph with set of vertices C and two vertices (i, λ) and (j, μ) adjacent if and only if $i = j$ or $\lambda = \mu$. In particular given $S = \mathcal{M}^0[G; I, \Lambda; P]$ we define $\mathbb{H}_S \subseteq I \times \Lambda$ as the set of coordinates of the group $\mathcal{H}$-classes of S, that is

$$\mathbb{H}_S = \{(i, \lambda) \in I \times \Lambda : H_{i\lambda} \text{ is a group}\} = \{(i, \lambda) \in I \times \Lambda : p_{\lambda i} \neq 0\}.$$

We will show, in Lemma 2.24, that the graph $\Gamma(\mathbb{H}_S)$ is connected if and only if the graph $\Delta(P)$ is connected. We say that S is *connected* if and only if $\Gamma(\mathbb{H}_S)$ (or equivalently $\Delta(P)$) is connected (see Figure 2.1 for examples).

Also, for $I' \subseteq I$ and $\Lambda' \subseteq \Lambda$ we say that $I' \times \Lambda'$ is a connected component of S precisely when the subgraph of $\Gamma(\mathbb{H}_S)$ induced by the vertices $I' \cup \Lambda'$ is a connected component of $\Gamma(\mathbb{H}_S)$.

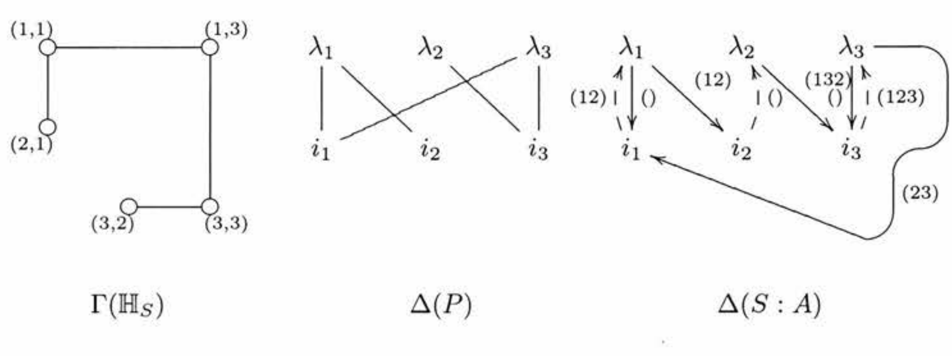
Example 2.19. Let $G = S_3 = \{(), (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$, the symmetric group of degree three. Define $S = \mathcal{M}^0[G; \{i_1, i_2, i_3\}, \{\lambda_1, \lambda_2, \lambda_3\}; P]$ where

$$P = \begin{pmatrix} () & (1\,2) & 0 \\ 0 & 0 & (1\ 3\ 2) \\ (2\ 3) & 0 & () \end{pmatrix}.$$

Let

$$A = \{(1, (1\ 2), 1), (2, (), 2), (3, (1\ 2\ 3), 3)\} \subseteq S.$$

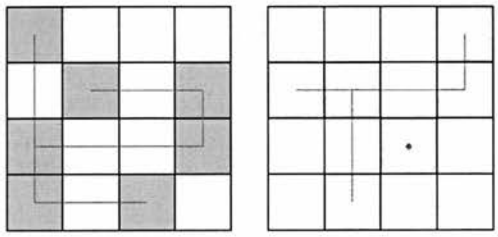
Then, for this example, the three graphs defined above are given below.



Since $\Gamma(\mathbb{H}_S)$ is connected, the semigroup S is a connected completely 0-simple semigroup.

Since $F(S) = \langle E(S) \rangle$ is a subsemigroup of S and $H_{i\lambda}$ is a subgroup of S the intersection $F(S) \cap H_{i\lambda}$ is a subsemigroup of $H_{i\lambda}$ in S and therefore must be a

Figure 2.1: Two egg-box pictures of $\mathcal{D}$-classes of completely 0-simple semigroups. The shaded boxes are the group $\mathcal{H}$-classes.



Connected Disconnected

subgroup of S. But what does this group look like? Clearly

$$F(S) \cap H_{i\lambda} \subseteq \{(i, k, \lambda) : k \in K\}$$

where K is the subgroup of G generated by the non-zero entries of the matrix P. In general, however, these two sets are not going to be equal.

Example 2.20. Let G be the cyclic group of order 5 written multiplicatively and generated by a: so $G = \{a^0, a^1, a^2, a^3, a^4\}$. Let $S = \mathcal{M}^0[G; \{1,2\}, \{1,2\}; P]$ with:

$$P = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}.$$

From the discussion above we deduce that:

$$F(S) \cap H_{11} \leq 1 \times G \times 1 = H_{11}.$$

However, by Theorem 2.54, the semigroup S is isomorphic to

$$T = \mathcal{M}^0[G; \{1,2\}, \{1,2\}; Q]$$

where:

$$Q = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

and hence for every group $\mathcal{H}$-class $H_{i\lambda}$ of $T \cong S$ we have:

$$F(T) \cap H_{i\lambda} = \{(1, a^0, 1)\}.$$

Since $T \cong S$ it follows that:

$$F(S) \cap H_{11} = \{(1, a^0, 1)\} \neq 1 \times G \times 1.$$

This example is important for the following reason. If we are interested in determining $\langle E(S) \rangle$ for S a finite 0-simple semigroup then, in the example above, the latter of the two representations is a more useful one. It satisfies the property that the subsemigroup generated by the idempotents intersected with a group $\mathcal{H}$-class is isomorphic to the subgroup of G generated by the non-zero entries in the matrix P. In fact, such a "nice" normalization always exists. This is called *Graham normal form* (see [41]) and will be discussed in detail in Section 2.10.

Returning to the problem of describing the group $F(S) \cap H_{i\lambda}$ we now show how this group relates to a group of paths in the graph $\Delta(P)$. Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be connected and let $(1_I, 1_\Lambda) \in I \times \Lambda$ with $p_{1_\Lambda 1_I} \neq 0$ so that $H_{1_I 1_\Lambda}$ is a group $\mathcal{H}$-class of S.

Lemma 2.21. *[81, Lemma 4.3] The mapping $\psi : H_{1_I 1_\Lambda} \to G$ defined by*

$$\psi((1_I, g, 1_\Lambda)) = g p_{1_\Lambda 1_I}$$

is a group isomorphism. It maps $H_{1_I 1_\Lambda} \cap F(S)$ onto $V_{1_I 1_\Lambda} p_{1_\Lambda 1_I}$.

Proof. It is routine to check that the map is an isomorphism. The second assertion follows from Theorem 2.18. □

We can actually say a lot more about the subgroups $V_{i\lambda} p_{\lambda i}$ of G.

Lemma 2.22. *Let $H_{i\lambda}$ and $H_{j\mu}$ be group $\mathcal{H}$-classes of a connected completely 0-simple semigroup $S = \mathcal{M}^0[G; I, \Lambda; P]$. Then the subgroups $V_{i\lambda} p_{\lambda i}$ and $V_{j\mu} p_{\mu j}$ are conjugate in G.*

Proof. Since S is connected we can fix a path π in $\Delta(S : E(S))$ from μ to i. We claim that with $g = p_{\mu j}^{-1} V(\pi)$ we have

$$g V_{i\lambda} p_{\lambda i} g^{-1} = V_{j\mu} p_{\mu j}.$$

It is sufficient to show:

$$p_{\mu j}^{-1} V(\pi) V_{i\lambda} p_{\lambda i} V(\pi)^{-1} \subseteq V_{j\mu}.$$

Let $p \in P_{i\lambda}$. In the graph $\Delta(S : E(S))$ this path may be extended to

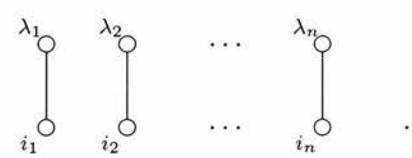
$$j \xrightarrow{p_{\mu j}^{-1}} \mu \xrightarrow[\pi]{\qquad} i \xrightarrow[p]{\qquad} \lambda \xrightarrow{p_{\lambda i}} i \xrightarrow[\pi^{-1}]{\qquad} \mu \in P_{j\mu}.$$

It follows that

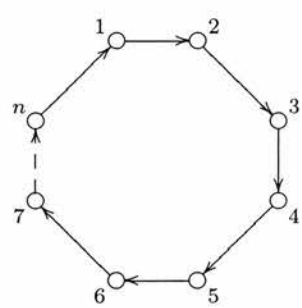
$$p_{\mu j}^{-1} V(\pi) V(p) p_{\lambda i} V(\pi)^{-1} \in V_{j\mu}.$$

Since this is true for every $p \in P_{i\lambda}$ the result follows. □

As promised earlier we now describe all the bases of the aperiodic Brandt semigroup B_n. The structure matrix P of B_n is the $n \times n$ identity matrix I_n. Therefore the graph $\Delta(P)$ has the following form.



Let $A \subseteq B_n$. Let $G(B_n, A)$ be the graph given by contracting the edges (λ_k, i_k), for all $1 \leq k \leq n$, in the graph $\Delta(B_n : A)$. Thus the graph $G(B_n, A)$ is isomorphic to the graph with vertex set $\{1, \ldots, n\}$ and set of directed edges equal to A. It follows from Lemma 2.17 that A generates B_n if and only if the graph $G(B_n, A)$ is strongly connected (i.e. there is a directed path between every pair of vertices). It follows from [8, Corollary 7.2.3] that if $|A| = n$ then $G(B_n, A)$ is strongly connected if and only if it is isomorphic to the following graph.



This proves the following result.

Proposition 2.23. *Let $S = B_n$ be the $n \times n$ aperiodic Brandt semigroup and let $\alpha \in S_n$. Then*

$$A_\alpha = \{(i, i\alpha) : 1 \leq i \leq n\}$$

generates S if and only if α is a n-cycle in S_n.

Figure 2.2: Diagram showing that the groups $V_{j\mu}p_{\mu j}$ and $V_{i\lambda}p_{\lambda i}$ are conjugate in G.

2.5 Connected completely 0-simple semigroups

The question of finding a formula for rank(S) divides into two cases: the case when S is connected and the case when it is not. The connected case was dealt with in [81] and in what follows we will extend these results to deal with the disconnected case. For what remains of this section S will denote a finite connected completely 0-simple semigroup. We now give the details of a number of results from [81] which will be needed later on.

Lemma 2.24. *[81, Theorem 2.1] The following conditions are equivalent for any completely* 0*-simple semigroup S:*

(i) $\Gamma(\mathbb{H}_S)$ *is a connected graph;*

(ii) $\Delta(P)$ *is a connected graph;*

(iii) $F(S) \cap H_{i\lambda} \neq \varnothing$ *for any* $i \in I$ *and any* $\lambda \in \Lambda$.

Proof. ((i) $\Leftrightarrow$ (ii)) The graph $\Gamma(\mathbb{H}_S)$ is disconnected if and only if there exist vertices (i, λ) and (j, μ) in different connected components of $\Gamma(\mathbb{H}_S)$ which is true if and only if the edges $\{i, \lambda\}$ and $\{j, \mu\}$ are in different connected components of the graph $\Delta(P)$. Such a pair of edges exists if and only if $\Delta(P)$ is not connected. ((ii) $\Rightarrow$ (iii)) Since the graph $\Gamma(\mathbb{H}_S)$ is connected it follows that the graph $\Delta(P)$ is connected. Given some $i \in I$ and $\lambda \in \Lambda$ let $p = (e_1, e_2, \ldots, e_k)$ be a directed path in $\Delta(P)$ starting at i and ending at λ. This is a valid path, as defined in the previous section, and the corresponding element $W(p) \in F(S)$ belongs to the $\mathcal{H}$-class $H_{i\lambda}$. Since i and λ were arbitrary it follows that $F(S) \cap H_{i\lambda} \neq \varnothing$ for all $(i, \lambda) \in I \times \Lambda$. ((iii) $\Rightarrow$ (i)) Suppose that $H_{i\lambda} \cap F(S)$ is empty. Then it follows from the definition of the graph that there is no path from i to λ in the graph $\Delta(P)$. It follows that $\Delta(P)$ is not connected and so neither is $\Gamma(\mathbb{H}_S)$.

□

In the next lemma we introduce the function $\phi(i, \lambda, j, \mu)$. This function will play a crucial role in what follows. The family of functions $\phi(i, \lambda, j, \mu)$ where

$i, j \in I$ and $\lambda, \mu \in \Lambda$ allow us to use the idempotents of connected completely 0-simple semigroups to "move between the $\mathcal{H}$-classes". More than this, when moving between two group $\mathcal{H}$-classes we are able to map in an isomorphic way.

Lemma 2.25. *[81, Theorem 2.1] Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a connected completely 0-simple semigroup. For any $i, j \in I$ and any $\lambda, \mu \in \Lambda$ there exist $p(i, \lambda, j, \mu), q(i, \lambda, j, \mu) \in F(S)$ such that the mapping $\phi(i, \lambda, j, \mu) : H_{i\lambda} \to H_{j\mu}$ defined by*

$$\phi(i, \lambda, j, \mu)(x) = p(i, \lambda, j, \mu)xq(i, \lambda, j, \mu)$$

is a bijection.

The elements $p(i, \lambda, j, \mu), q(i, \lambda, j, \mu)$ can be chosen so that

$$\phi(i, \lambda, j, \mu)^{-1} = \phi(j, \mu, i, \lambda)$$

and $\phi(i, \lambda, j, \mu)$ is a group isomorphism if both $H_{i\lambda}$ and $H_{j\mu}$ are groups.

Proof. First we show that a bijection can be found with the given properties. Then we show that in the special case where $H_{i\lambda}$ and $H_{j\mu}$ are both groups this bijection may be chosen to be an isomorphism.

In the graph $\Delta(P)$ choose and fix a path π which starts at j and finishes at i. Such a path exists since the graph $\Delta(P)$ is connected. Say this path is:

$$j \to \mu_1 \to j_2 \to \mu_2 \to \ldots \to j_m \to \mu_m \to i.$$

This is not a valid path but if we shorten it slightly we do get a valid path. Let π' denote the path π but restricted from j to μ_m. In a similar way let ξ be a fixed path in the graph $\Delta(P)$ connecting λ to μ. Such a path exists since the graph is connected. Say the path ξ is:

$$\lambda \to i_1 \to \lambda_1 \to i_2 \to \ldots \to \lambda_n \to i_n \to \mu.$$

Once again, this path is not valid in $\Delta(P)$ but if we shorten it, by removing the initial vertex, we get a path starting at i_1 and ending at μ which is valid. Call this path ξ'. By the definition of $\Delta(P)$ we have $W(\pi') \in F(S)$ and $W(\xi') \in F(S)$.

Now define $p(i, \lambda, j, \mu) = W(\pi')$ and $q(i, \lambda, j, \mu) = W(\xi')$. The map

$$\phi(i, \lambda, j, \mu)(x) = p(i, \lambda, j, \mu)xq(i, \lambda, j, \mu)$$

is a bijection from $H_{i\lambda}$ to $H_{j\mu}$ since it sends $(i, g, \lambda) \in H_{i\lambda}$ to $(j, V(\pi)gV(\xi), \mu)$ where $V(\pi), V(\xi) \in G$ are both fixed.

Now define the path π^{-1} from i to j to be the reverse of the path π and the path ξ^{-1} to be the reverse of the path ξ. The paths $\pi^{-1'}$ and $\xi^{-1'}$ are defined in the analogous way to above and then define $p(j,\mu,i,\lambda) = W(\pi^{-1'})$ and $q(j,\mu,i,\lambda) = W(\xi^{-1'})$. In this way the map $\phi(j,\mu,i,\lambda)$ maps $(j,h,\mu) \in H_{j\mu}$ to $(i, V(\pi)^{-1}hV(\xi)^{-1}, \lambda)$. Therefore for every $(i,g,\lambda) \in H_{i\lambda}$ we have:

$$\begin{aligned}\phi(j,\mu,i,\lambda)(\phi(i,\lambda,j,\mu)((i,g,\lambda))) &= \phi(j,\mu,i,\lambda)((j,V(\pi)gV(\xi),\mu)) \\ &= (i, V(\pi)^{-1}V(\pi)gV(\xi)V(\xi)^{-1},\lambda) \\ &= (i,g,\lambda)\end{aligned}$$

and so $\phi(i,\lambda,j,\mu)^{-1} = \phi(j,\mu,i,\lambda)$.

Now consider the special case where both $H_{i\lambda}$ and $H_{j\mu}$ are group $\mathcal{H}$-classes. This means that the edges $i \xrightarrow{p_{\lambda i}^{-1}} \lambda$ and $j \xrightarrow{p_{\mu j}^{-1}} \mu$ belong to the graph $\Delta(P)$. This gives some control over the choice of the paths π and ξ in the above construction. Carry out the same process as above but this time once π has been fixed we may define

$$\xi = \lambda \xrightarrow{p_{\lambda i}} i \xrightarrow{\pi^{-1}} j \xrightarrow{p_{\mu j}^{-1}} \mu.$$

Now the map $\phi = \phi(i,\lambda,j,\mu)$ maps $(i,g,\lambda) \in H_{i\lambda}$ to $(j, V(\pi)gp_{\lambda i}V(\pi)^{-1}p_{\mu j}^{-1}, \mu) \in H_{j\mu}$. Again, this map is clearly a bijection. Moreover, it is an isomorphism since:

$$\begin{aligned}\phi((i,g,\lambda)(i,h,\lambda)) &= \phi(i, gp_{\lambda i}h, \lambda) \\ &= (j, V(\pi)(gp_{\lambda i}h)p_{\lambda i}V(\pi)^{-1}p_{\mu j}^{-1}, \mu) \\ &= (j, V(\pi)(gp_{\lambda i})(hp_{\lambda i})V(\pi)^{-1}p_{\mu j}^{-1}, \mu) \\ &= (j, V(\pi)(gp_{\lambda i})V(\pi)^{-1}p_{\mu j}^{-1}p_{\mu j}V(\pi)(hp_{\lambda i})V(\pi)^{-1}p_{\mu j}^{-1}, \mu) \\ &= (j, V(\pi)gp_{\lambda i}V(\pi)^{-1}p_{\mu j}^{-1}, \mu)(j, V(\pi)hp_{\lambda i}V(\pi)^{-1}p_{\mu j}^{-1}, \mu) \\ &= \phi((i,g,\lambda))\phi((i,h,\lambda)).\end{aligned}$$

□

Lemma 2.26. *Let $H_{i\lambda}$ and $H_{j\mu}$ be group $\mathcal{H}$-classes of the connected completely 0-simple semigroup S. There is an isomorphism $\theta : H_{i\lambda} \to H_{j\mu}$ such that $(F(S) \cap H_{i\lambda})\theta = F(S) \cap H_{j\mu}$.*

Proof. The map $\phi = \phi(i,\lambda,j,\mu) : H_{i\lambda} \to H_{j\mu}$ by Lemma 2.25 is an isomorphism. We claim that $\phi(F(S) \cap H_{i\lambda}) = F(S) \cap H_{j\mu}$. Indeed, if $x \in F(S) \cap H_{i\lambda}$ then $\phi(x) \in F(S)xF(S) \in F(S)^3 = F(S)$ and so $\phi(F(S) \cap H_{i\lambda}) \subseteq F(S) \cap H_{j\mu}$. Since ϕ is a bijection and since, by Lemma 2.22 the groups $F(S) \cap H_{i\lambda}$ and $F(S) \cap H_{j\mu}$

are isomorphic, it follows that $\phi(F(S) \cap H_{i\lambda}) = F(S) \cap H_{j\mu}$. □

So the subgroups $F(S) \cap H_{i\lambda}$ and $F(S) \cap H_{j\mu}$ are not only isomorphic but they sit inside their respective group $\mathcal{H}$-classes in the same way. The following result now follows.

Corollary 2.27. *Let S be a connected completely 0-simple semigroup with $H_{i\lambda}$ and $H_{j\mu}$ two group $\mathcal{H}$-classes of S. Then*

$$\text{rank}(H_{i\lambda} : F(S) \cap H_{i\lambda}) = \text{rank}(H_{j\mu} : F(S) \cap H_{j\mu}).$$

The next lemma, roughly speaking, shows us how effective a given subset A of S can be in helping to generate the elements of a fixed group $\mathcal{H}$-class $H_{i\lambda}$.

Lemma 2.28. *[81, Lemma 3.4] Let $A = \{a_1, \ldots, a_r\} \subseteq S$ where $a_j \in H_{i_j\lambda_j}, j = 1, \ldots, r$ and let $H_{i\lambda}$ be a group $\mathcal{H}$-class. If we write*

$$B = \{\phi(i_1, \lambda_1, i, \lambda)(a_1), \ldots, \phi(i_r, \lambda_r, i, \lambda)(a_r)\} \subseteq H_{i\lambda}$$

then

$$\langle F(S) \cup A\rangle \cap H_{i\lambda} = \langle (F(S) \cap H_{i\lambda}) \cup B\rangle.$$

Proof. ($\supseteq$) First note that by definition $B \subseteq F(S)AF(S)$ and so $B = B \cap H_{i\lambda} \subseteq F(S)AF(S) \cap H_{i\lambda}$. It follows that:

$$\begin{aligned} \langle (F(S) \cap H_{i\lambda}) \cup B\rangle &\subseteq \langle (F(S) \cap H_{i\lambda}) \cup (F(S)AF(S) \cap H_{i\lambda})\rangle \\ &\subseteq \langle F(S) \cup F(S)AF(S)\rangle \cap H_{i\lambda} \\ &\subseteq \langle F(S) \cup A\rangle \cap H_{i\lambda}. \end{aligned}$$

($\subseteq$) First observe that $A \subseteq F(S)BF(S)$ since for $a \in A \cap H_{j\mu}$ we have

$$\begin{aligned} a &= \phi(i, \lambda, j, \mu)(\phi(j, \mu, i, \lambda)(a)) = p(i, \lambda, j, \mu)\phi(j, \mu, i, \lambda)(a)q(i, \lambda, j, \mu) \\ &\in F(S)\phi(j, \mu, i, \lambda)(a)F(S) \subseteq F(S)BF(S). \end{aligned}$$

It follows that:

$$\begin{aligned} \langle F(S) \cup A\rangle \cap H_{i\lambda} &\subseteq \langle F(S) \cup F(S)BF(S)\rangle \cap H_{i\lambda} \\ &\subseteq \langle F(S) \cup B\rangle \cap H_{i\lambda}. \end{aligned}$$

Also:

$$\langle F(S) \cup B\rangle \cap H_{i\lambda} = \langle E(S) \cup B\rangle \cap H_{i\lambda} \subseteq \langle (F(S) \cap H_{i\lambda}) \cup B\rangle.$$

Indeed, any product in $\langle E(S) \cup B \rangle \cap H_{i\lambda}$ has the form:

$$e_1 \dots e_{r_1} b_1 e_{r_1+1} \dots e_{r_2} b_2 e_{r_2+1} \dots e_{r_{k-1}} b_{k-1} e_{r_{k-1}+1} \dots e_{r_k}$$

where each $b_m \in H_{i\lambda}$ and $e_n \in E(S)$ for all n. Let $e_{i\lambda}$ denote the idempotent of $H_{i\lambda}$ which is the identity of this group $\mathcal{H}$-class. The above product is equal to:

$$(e_1 \dots e_{r_1} e_{i\lambda}) b_1 (e_{i\lambda} e_{r_1+1} \dots e_{r_2} e_{i\lambda}) b_2 (e_{i\lambda} e_{r_2+1} \dots e_{r_{k-1}} e_{i\lambda}) b_{k-1} (e_{i\lambda} e_{r_{k-1}+1} \dots e_{r_k})$$

which has the form

$$f_1 b_1 f_2 b_2 \dots b_{k-1} f_k$$

where $f_l \in F(S) \cap H_{i\lambda}$ and $b_l \in B$ for all l. The reason that f_1 and f_k belong to $H_{i\lambda}$ is because the entire product $f_1 b_1 f_2 b_2 \dots b_{k-1} f_k$ belongs to $H_{i\lambda}$. It follows that:

$$\langle F(S) \cup A \rangle \cap H_{i\lambda} \subseteq \langle F(S) \cup B \rangle \cap H_{i\lambda} \subseteq \langle (F(S) \cap H_{i\lambda}) \cup B \rangle$$

as required. □

Using the previous result we may obtain a result concerning the relative rank of any subset of $F(S)$ in S.

Lemma 2.29. *Let S be a connected completely 0-simple semigroup with $H_{i\lambda}$ a group $\mathcal{H}$-class and $U \subseteq F(S)$. Then*

$$\operatorname{rank}(S : U) \geq \operatorname{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S)).$$

Proof. Let $V \subseteq S$ be such that $\langle U \cup V \rangle = S$, say $V = \{v_1, \dots, v_m\} \subseteq S$ where $v_k \in H_{i_k \lambda_k}$ $(k = 1, \dots, m)$. If we write

$$B = \{\phi(i_1, \lambda_1, i, \lambda)(v_1), \dots, \phi(i_m, \lambda_m, i, \lambda)(v_m)\} \subseteq H_{i\lambda}$$

then by Lemma 2.28 we have

$$\langle F(S) \cup V \rangle \cap H_{i\lambda} = \langle (F(S) \cap H_{i\lambda}) \cup B \rangle.$$

Now since $U \cup V$ is a generating set for S, we have

$$\begin{aligned} H_{i\lambda} &= S \cap H_{i\lambda} = \langle U \cup V \rangle \cap H_{i\lambda} = \langle F(S) \cup U \cup V \rangle \cap H_{i\lambda} \\ &= \langle F(S) \cup V \rangle \cap H_{i\lambda} = \langle (F(S) \cap H_{i\lambda}) \cup B \rangle. \end{aligned}$$

Therefore

$$|V| \geq |B| \geq \operatorname{rank}(H_{i\lambda} : F(S) \cap H_{i\lambda}).$$

□

This leads to the following corollary which gives an important lower bound for ranks of connected completely 0-simple semigroups.

Corollary 2.30. *[81, Lemma 3.6] If $H_{i\lambda}$ is a group $\mathcal{H}$-class then*

$$\operatorname{rank}(S) \geq \operatorname{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S)).$$

Proof. By Lemma 2.29 we have:

$$\operatorname{rank}(S) = \operatorname{rank}(S : \varnothing) \geq \operatorname{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S))$$

as required. □

2.6 Arbitrary completely 0-simple semigroups

Now we look at the case when $\Gamma(\mathbb{H}_S)$ is not necessarily connected. For the remainder of this chapter, unless otherwise stated, S will denote a completely 0-simple semigroup, represented as a Rees matrix semigroup $\mathcal{M}^0[G; I, \Lambda; P]$, with k connected components $I_1 \times \Lambda_1, \ldots, I_k \times \Lambda_k$ so the matrix P has the form suggested by the following picture:

$$\begin{array}{c} \\ \Lambda_1 \\ \Lambda_2 \\ \vdots \\ \Lambda_k \end{array} \begin{array}{c} \begin{array}{cccc} I_1 & I_2 & \ldots & I_k \end{array} \\ \begin{pmatrix} C_1 & & & 0 \\ & C_2 & & \\ & & \ddots & \\ 0 & & & C_k \end{pmatrix} \end{array}.$$

We will first consider some properties that all generating sets of S must have.

Definition 2.31. We say that $A \subseteq S$ is an *$\mathcal{H}$-class transversal generating set* if

$$\langle A \rangle \cap H_{i\lambda} \neq \varnothing \text{ for all } i \in I \text{ and } \lambda \in \Lambda.$$

For example, Theorem 2.24 tells us that $\Gamma(\mathbb{H}_S)$ is connected if and only if $E(S)$ is an $\mathcal{H}$-class transversal generating set. Clearly the smallest size such a set can have is $\max(|I|, |\Lambda|)$. In fact, as a consequence of the results of Section 2.3, concerning the rank of an arbitrary rectangular 0-band, we know that we can always find at least one $\mathcal{H}$-class transversal generating set with this minimal size.

Lemma 2.32. *There exists $A \subseteq S$ such that $|A| = \max(|I|, |\Lambda|)$ and A is an $\mathcal{H}$-class transversal generating set.*

Proof. The assertion follows from Theorem 2.10 and from the fact that the mapping $(i, g, \lambda) \mapsto (i, \lambda)$ defines an epimorphism from S onto a rectangular 0-band. □

An $\mathcal{H}$-class transversal generating set with size $\max(|I|, |\Lambda|)$ will be the first building block we will use when constructing minimal generating sets for S. We will call an $\mathcal{H}$-class transversal generating set with size $\max(|I|, |\Lambda|)$ an *$\mathcal{H}$-class transversal basis.*

Definition 2.33. We call $C \subseteq I \times \Lambda$ *component connecting coordinates* if $\Gamma(\mathbb{H}_S \cup C)$ is connected. Similarly we will call $A \subseteq S$ a *component connecting set* if $\{(i, \lambda) | (i, g, \lambda) \in A\} \subseteq I \times \Lambda$ is a set of component connecting coordinates.

If $\Gamma(\mathbb{H}_S)$ has k connected components then the smallest size a component connecting set can have is $k - 1$. We will call these the *minimal connecting sets.*

Lemma 2.34. *Every component connecting set D has a subset $E \subseteq D$ that is minimal.*

Proof. Let A be the coordinates of the component connecting set D. Let $C_1, \ldots, C_k$ be the connected components of $\Gamma(\mathbb{H}_S)$. Construct a new graph Γ' with vertices $C_1, \ldots, C_k$ and C_i adjacent to C_j if and only if there is some $c_i \in C_i$, $c_j \in C_j$ and $a \in A$ such that $\{c_i, a\}$ and $\{a, c_j\}$ are edges in $\Gamma(\mathbb{H}_S)$. Then Γ' is a connected graph with k vertices and so has a spanning tree with $k - 1$ edges. This spanning tree corresponds, in an obvious way, to a minimal set of coordinates $A' \subseteq A$ which in turn correspond to a subset $E \subseteq D$ that is minimal. □

Lemma 2.35. *Every $\mathcal{H}$-class transversal generating set is component connecting.*

Proof. Suppose otherwise. Let A be an $\mathcal{H}$-class transversal generating set that is not component connecting. Then we can choose $(i, \lambda), (j, \mu) \in I \times \Lambda$ such that (i, λ) and (j, μ) are in different connected components of $\Gamma(\mathbb{H}_S \cup A)$. It follows that i and μ are in different connected components of the graph $\Delta(S : A)$ and so by Lemma 2.17 we have $\langle A \rangle \cap H_{i\mu} = \varnothing$, a contradiction. □

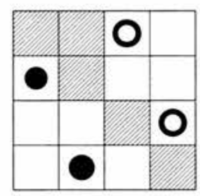
Corollary 2.36. *Every $\mathcal{H}$-class transversal generating set (and in particular every $\mathcal{H}$-class transversal basis) has a component connecting subset which is minimal.*

Example 2.37. Let $G = \{1, a\}$ be the cyclic group of order 2 and let

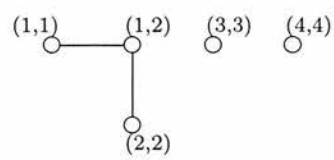
$$S = \mathcal{M}^0[G; \{1, \dots, 4\}, \{1, \dots, 4\}; P]$$

where

$$P = \begin{pmatrix} a & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad S \cong$$



.



$\Gamma(\mathbb{H}_S)$

The graph $\Delta(P)$ has 3 connected components.

(i) $C = \{(1,3), (2,1), (3,4), (4,2)\} \subseteq I \times \Lambda$ is a set of component connecting coordinates.

(ii) $A = \{(1,a,3), (2,1,1), (3,1,4), (4,a,2)\} \subseteq S$ is an $\mathcal{H}$-class transversal basis which is a component connecting set.

(iii) $T = \{(1,a,3), (3,1,4)\} \subseteq A$ is a minimal component connecting subset of A.

We now define $r_{\min}$ which is the most complicated term that will appear in the formula for the rank of a 0-Rees matrix semigroup. We use $\mathrm{Map}(X,Y)$ to denote the set of all maps from a set X into a set Y.

Definition 2.38. Given the structure matrix P of S we define $P_{C,\theta}$ where $C \subseteq \Lambda \times I$ and $\theta \in \mathrm{Map}(C, G)$ to be a new $|\Lambda| \times |I|$ matrix with entries $p^*_{\lambda i}$ where

$$p^*_{\lambda i} = \begin{cases} (\lambda, i)\theta & \text{if } (\lambda, i) \in C \\ p_{\lambda i} & \text{otherwise.} \end{cases}$$

Also let $S_{C,\theta} = \mathcal{M}^0[G; I, \Lambda; P_{C,\theta}]$. When $|C| = 1$ we will use the more relaxed notation $S(\lambda, i, (\lambda, i)\theta)$.

Since, depending on the circumstances, we may need to view sets of coordinates sometimes as subsets of $I \times \Lambda$ and at other times as subsets of $\Lambda \times I$ we make the following definition.

Definition 2.39. Given $A \subseteq I \times \Lambda$ we define A^T as

$$A^T = \{(\lambda, i) : (i, \lambda) \in A\} \subseteq \Lambda \times I,$$

and call this set the *transpose* of A.

Definition 2.40. Let $H_{i\lambda}$ be any group $\mathcal{H}$-class of S and let $C \subseteq I \times \Lambda$ be a set of component connecting coordinates with minimal size. Then define

$$r_{\min} = \min_{\theta \in \mathrm{Map}(C^T, G)} (\mathrm{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S_{C^T,\theta}))).$$

If S is connected then $C = \varnothing$ and $r_{\min} = \mathrm{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S))$. Otherwise S is not connected and we associate the family $\{S_{C^T,\theta} : \theta \in \mathrm{Map}(C^T, G)\}$ of connected completely 0-simple semigroups with S. We then look through this family searching for a member that minimizes the relative rank of $F(S)$ in a group $\mathcal{H}$-class.

We have to show $r_{\min}$ is well defined i.e. that it does not depend on the choice of $H_{i\lambda}$ or on the choice of C. First we note that, since $S_{C^T,\theta}$ is connected, by Corollary 2.27 the number $r_{\min}$ is independent of the choice of $H_{i\lambda}$. That $r_{\min}$ does not depend on the choice of C will eventually be proved in Lemma 2.43.

Lemma 2.41. *Suppose that P has two connected components $\Lambda_1 \times I_1$ and $\Lambda_2 \times I_2$, and let $H_{i\lambda}$ be a group $\mathcal{H}$-class. If $(\mu, j), (\nu, k) \in (\Lambda_1 \times I_2) \cup (\Lambda_2 \times I_1)$ then for every $g \in G$ there exists $g' \in G$ such that*

$$H_{i\lambda} \cap F(S(\mu, j, g)) = H_{i\lambda} \cap F(S(\nu, k, g')).$$

Proof. There are essentially two cases to consider.

Case 1: $(\mu, j), (\nu, k) \in \Lambda_1 \times I_2$. Let $(i, \lambda) \in I \times \Lambda$ be arbitrary. We will show that given $(\mu, j), (\nu, k)$ and $g \in G$ we can choose $g' \in G$ so that if we let $S_1 = S(\mu, j, g)$ with structure matrix P_1 and $S_2 = S(\nu, k, g')$ with matrix P_2, then we have

$$V_{i\lambda}^{S_1} = V_{i\lambda}^{S_2}.$$

Here $V_{i\lambda}^{U}$ denotes the set of values of all paths from i to λ (so far denoted simply as $V_{i\lambda}$) in the graph $\Delta(Q)$ where Q is the structure matrix of U. This convention will be used throughout and will also apply to sets of paths $P_{i\lambda}$. Let us choose and fix a path $\pi_{\nu\mu}$ in $P^S_{\nu\mu}$ and define $w_{\nu\mu} = V(\pi_{\nu\mu}) \in V^S_{\nu\mu}$. Also, choose and fix a path π_{jk} in P^S_{jk} defining $w_{jk} = V(\pi_{jk}) \in V^S_{jk}$. Such paths exist since $I_1 \cup \Lambda_1$ and $I_2 \cup \Lambda_2$ are both connected components in $\Delta(P)$. Now define

$$g' = w_{\nu\mu} g w_{jk} \in V^S_{\nu\mu} g V^S_{jk}$$

so that

$$g = w^{-1}_{\nu\mu} g' w^{-1}_{jk} \in V^S_{\mu\nu} g' V^S_{kj}.$$

We observe that $\Delta(P_1)$ is connected and is precisely $\Delta(P)$ with the extra edge $\mu \xrightarrow{g} j$. Also $\Delta(P_2)$ is connected and is precisely $\Delta(P)$ with the extra edge $\nu \xrightarrow{g'} k$. Let p be an arbitrary path from i to λ in $\Delta(P_1)$. We show that there is a corresponding path p^* in $\Delta(P_2)$ with the same value. While following the path p whenever we come across

$$\ldots \to \mu \to j \to \ldots$$

we replace it with

$$\ldots \to \mu \to \nu \to k \to j \to \ldots.$$

Recall that the graph $\Delta(P)$ is directed and has unlabelled edges. Also, paths are uniquely determined by giving a list of vertices. Above we move from μ to ν using the fixed path $\pi_{\mu\nu}$ and from k to j using the path π_{kj}. Clearly $V(p^*) = V(p)$ by definition of g' and so $V^{S_1}_{i\lambda} \subseteq V^{S_2}_{i\lambda}$. Similarly $V^{S_2}_{i\lambda} \subseteq V^{S_1}_{i\lambda}$ and the result follows from Theorem 2.18. The case where $(\mu, j), (\nu, k) \in \Lambda_2 \times I_1$ is dual to this case.

Case 2: $(\mu, j) \in \Lambda_1 \times I_2$ *and* $(\nu, k) \in \Lambda_2 \times I_1$. We follow a very similar argument but this time let $w_{\mu j} \in V^S_{\mu j}$ and $w_{\nu k} \in V^S_{\nu k}$, which is possible since $I_1 \cup \Lambda_1$ and $I_2 \cup \Lambda_2$ are both connected components in $\Delta(P)$. Then g' is chosen so that the paths

$$\mu \to k$$

and

$$\mu \longrightarrow j \to \nu \longrightarrow k$$

have the same value. □

Lemma 2.42. *Let $\mathcal{T}$ be the family of all spanning trees of the complete graph*

K_n. *If* Γ *denotes the graph with vertex set* $\mathcal{T}$ *and edges defined by*

$$(T_1, T_2) \in \mathcal{E}(\Gamma) \Leftrightarrow \exists e_1, e_2 \in \mathcal{E}(K_n) : (T_1 - \{e_1\}) \cup \{e_2\} = T_2$$

then Γ *is connected.*

Proof. Observe that the set of acyclic subsets of the edge set of a graph Γ defines a matroid (see [15, Chapter 12]). Then since in particular the spanning trees of K_n are acyclic edge sets they are independent sets in the corresponding matroid. That there is a path between the vertices $T_1, T_2 \in \mathcal{T}$ is now a consequence of the exchange axiom for matroids. □

Combining the previous two lemmas, we conclude the following which tells us that $r_{\min}$ is indeed well defined.

Lemma 2.43. *Let* $S = \mathcal{M}^0[G; I, \Lambda; P]$ *be a completely* 0*-simple semigroup with connected components* $I_1 \times \Lambda_1, \ldots, I_k \times \Lambda_k$. *Let* $C, D \subseteq I \times \Lambda$ *be two sets of minimal component connecting coordinates. If* $H_{i\lambda}$ *is a group* $\mathcal{H}$*-class in* S *then*

$$\min_{\theta \in \mathrm{Map}(C^T, G)} (\mathrm{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S_{C^T, \theta}))) = \min_{\phi \in \mathrm{Map}(D^T, G)} (\mathrm{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S_{D^T, \phi}))).$$

Proof. First we note that $\Gamma(\mathbb{H}_S \cup C)$ and $\Gamma(\mathbb{H}_S \cup D)$ are connected. We remove one $c \in C$, changing $\Gamma(\mathbb{H}_S \cup C)$ into $\Gamma(\mathbb{H}_S \cup (C \setminus \{c\}))$ which has two connected components. Now by Lemma 2.42 we can find $d \in D$ such that $\Gamma(\mathbb{H}_S \cup (C \setminus \{c\}) \cup \{d\})$ is connected and by Lemma 2.41 we can replace θ by θ' so that $\mathrm{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S_{C^T, \theta}))$ equals $\mathrm{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S_{((C \setminus \{c\}) \cup \{d\})^T, \theta'}))$. Repeating this process we can, by Lemma 2.42, move from C to D in a finite number of steps, and, by Lemma 2.41, keep the relative ranks equal at each step. □

2.7 Substitution lemma

In this section we analyse how an element $a \in S$ can compensate for a "missing" entry of the matrix P. We will see how generators may be used to "connect" S back together.

First we need some more notation. Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ and $T = \mathcal{M}^0[G; I, \Lambda; Q]$. Note that S and T are equal as sets. For $A \subseteq S = T$ we will write $\langle A \rangle_S$ and $\langle A \rangle_T$ to mean the subsemigroups generated by the set A in S and T respectively.

Lemma 2.44. *Let* $S = \mathcal{M}^0[G; I, \Lambda; P]$. *Let* $i, j \in I$ *and* $\lambda, \mu \in \Lambda$ *be such that* $H_{j\lambda}, H_{i\mu}$ *are group* $\mathcal{H}$*-classes and* $H_{j\mu}, H_{i\lambda}$ *are* not *group* $\mathcal{H}$*-classes. If* $A \subseteq S$

with $a = (i, g, \lambda) \in A$ *and* $T = S(\mu, j, p_{\mu i} g p_{\lambda j})$ *then*

$$\langle A \rangle_S = \langle A \rangle_T.$$

Proof. Let $\Delta_S = \Delta(S : A)$ and $\Delta_T = \Delta(T : A)$ with E_S and E_T the edge sets of Δ_S and Δ_T respectively. The only difference between E_S and E_T is that E_T has the extra edge $\mu \xrightarrow{p_{\mu i} g p_{\lambda j}} j$. Let R_S and R_T be the sets of valid paths in Δ_S, Δ_T respectively. It follows from Lemma 2.17 that

$$\langle A \rangle_S = \{W(p) : p \in R_S\} = \{(k, V(p), \xi) : k \in I, \xi \in \Lambda, p \in \mathcal{P}^S_{k,\xi}\}$$

and

$$\langle A \rangle_T = \{W(p) : p \in R_T\} = \{(k, V(p), \xi) : k \in I, \xi \in \Lambda, p \in \mathcal{P}^T_{k,\xi}\}.$$

It is obvious that $\langle A \rangle_S \subseteq \langle A \rangle_T$ since any valid path in Δ_S is a valid path in Δ_T and so $R_S \subseteq R_T$. For the converse we argue in much the same way as in Lemma 2.41. For $p \in R_T$ we show that there is a corresponding path $p^* \in R_S$ with $V(p) = V(p^*)$. In p whenever we come across

$$\ldots \to \mu \xrightarrow{p_{\mu i} g p_{\lambda j}} j \to \ldots$$

we replace it with

$$\ldots \to \mu \xrightarrow{p_{\mu i}} i \xrightarrow{g} \lambda \xrightarrow{p_{\lambda j}} j \to \ldots$$

Clearly $V(p^*) = V(p)$ and so $\langle A \rangle_S = \langle A \rangle_T$. □

We now, after first making another couple of technical definitions, combine the main ideas of this section together in the form of Corollaries 2.47 and 2.48 which will be used in the proof of the main theorem of this chapter in the next section.

Definition 2.45. Let $C \subseteq I \times \Lambda$ be a minimal set of component connecting coordinates for S. For each $(i, \lambda) \in C$ choose and fix a pair $(\psi(\lambda), \phi(i)) \subseteq I \times \Lambda$ such that $p_{\phi(i)i} \neq 0$ and $p_{\lambda\psi(\lambda)} \neq 0$. Define

$$C^\delta = \{(\psi(\lambda), \phi(i)) : (i, \lambda) \in C\}$$

noting that $C^\delta \subseteq I \times \Lambda$. We call C^δ the *dual of* C *determined by* ψ *and* ϕ, or more often just the *dual of* C (see Figure 2.3).

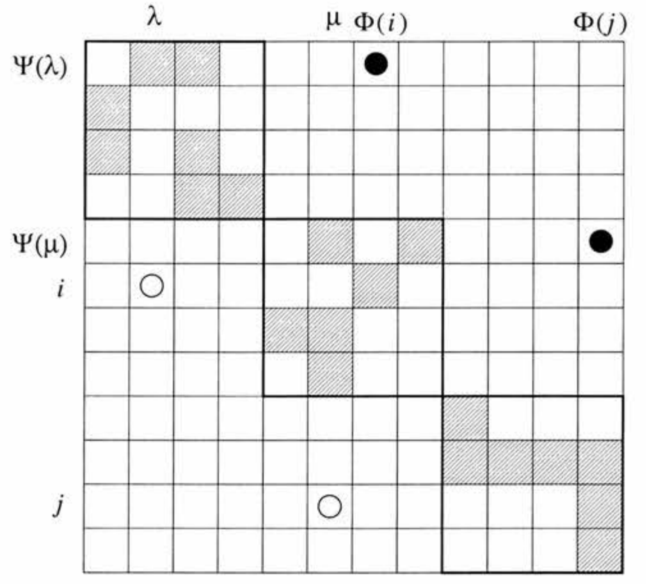


Figure 2.3: Egg-box picture of a 3 component completely 0-simple semigroup showing a set of minimal connecting coordinates (in white) and its dual (in black) (determined by Φ and Ψ).

Note that if C is a minimal set of component connecting coordinates then so is C^δ.

Definition 2.46. Given S as above and $B \subseteq S$ a minimal component connecting subset of S with coordinates C we define the connected completely 0-simple semigroup $\text{Con}(S, B)$ by $\text{Con}(S, B) = S_{(C^\delta)^T,\alpha}$ where C^δ is the dual of C and for $(i, g, \lambda) \in B$ we have $(\phi(i), \psi(\lambda))\alpha = p_{\phi(i)i}gp_{\lambda\psi(\lambda)}$. We call $\text{Con}(S, B)$ the *completion* of S by B.

Note that $\text{Con}(S, B)$ is a connected completely 0-simple semigroup.

Corollary 2.47. *For any set of minimal connecting coordinates C there is a set $X \subseteq S$ with coordinates C such that* $\text{rank}(H_{i\lambda} : H_{i\lambda} \cap F(\text{Con}(S, X))) = r_{\min}$.

Proof. Since $\text{Con}(S, X) = S_{(C^\delta)^T,\alpha}$ where for $(i, g, \lambda) \in X$ we have

$$(\phi(i), \psi(\lambda))\alpha = p_{\phi(i)i}gp_{\lambda\psi(\lambda)}$$

it follows that we can choose middle components of X so that α satisfies

$$\text{rank}(H_{i\lambda} : H_{i\lambda} \cap F(\text{Con}(S, X))) = \text{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S_{(C^\delta)^T,\alpha})) = r_{\min}.$$

□

Corollary 2.48. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ with connected components $I_1 \times \Lambda_1, \dots, I_k \times \Lambda_k$, let $B \subseteq S$ be a minimal component connecting set with coordinates $C \subseteq I \times \Lambda$, let $H_{i\lambda}$ be a group $\mathcal{H}$-class in S, let $T = \mathrm{Con}(S, B)$ and let $A \subseteq S$ be such that $B \subseteq A$. We have:*

(i) $B \subseteq F(T)$;

(ii) $\langle A \rangle_S = \langle A \rangle_T$;

(iii) $\mathrm{rank}(S : B) = \mathrm{rank}(T : B)$.

Proof. (i) If $(i, g, \lambda) \in B$ then $(i, g, \lambda) = (i, p_{i\phi(i)}^{-1}, \phi(i))(\psi(\lambda), p_{\psi(\lambda)\lambda}^{-1}, \lambda)$, a product of two idempotents. (ii) Repeated application of Lemma 2.44. (iii) A direct consequence of (ii). □

2.8 Main theorem

Given a generating set for $K = S\natural$, the natural rectangular 0-band homomorphic image of the finite completely 0-simple semigroup S, we will show how we can build a generating set for S around its coordinates.

Proposition 2.49. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be any completely 0-simple semigroup with connected components $I_1 \times \Lambda_1, \dots, I_k \times \Lambda_k$. Let K be the natural rectangular 0-band homomorphic image of S. Let B be a generating set for K. Then there exists a generating set X of S such that*

$$|X| = \max(|B|, r_{\min} + k - 1)$$

where

$$r_{\min} = \min_{\theta \in \mathrm{Map}(C^T, G)} (\mathrm{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S_{C^T, \theta}))),$$

and where $C \subseteq I \times \Lambda$ is any minimal set of component connecting coordinates, and $H_{i\lambda}$ is any group $\mathcal{H}$-class. In particular, if $\natural : S \to K$ is the natural homomorphism $(i, g, \lambda) \mapsto (i, \lambda)$ then X may be chosen to satisfy $B \subseteq X\natural$.

Proof. Let $B \subseteq I \times \Lambda$ be a generating set of K. Let D be a minimal connecting subset of B, which exists by Corollary 2.36, so $|D| = k - 1$. By Corollary 2.47 we can find $A \subseteq S$ with coordinates D such that for $T = \mathrm{Con}(S, A)$ we have $\mathrm{rank}(H_{i\lambda} : H_{i\lambda} \cap F(T)) = r_{\min}$. We will extend A to a generating set for T, which, by Corollary 2.48, will serve as a generating set for S also. Let $U \subseteq H_{i\lambda}$ be such that $|U| = r_{\min}$ and $\langle U \cup (H_{i\lambda} \cap F(T)) \rangle = H_{i\lambda}$, say $U = \{u_1, \dots, u_{r_{\min}}\}$. There are two cases to consider:

Case 1: $|B| \geq r_{\min} + k - 1$. First we note that

$$|B \setminus D| = |B| - (k-1) \geq r_{\min} + (k-1) - (k-1) = r_{\min}.$$

We can extend A to $X = A \cup Y \subseteq T$ with coordinates B in the following way. Let $Y = \{v_1, \ldots, v_{r_{\min}}, \ldots, v_{|B|-(k-1)}\} \subseteq T$ have coordinates corresponding to $B \setminus D$. Since T is connected we can choose middle components of $v_1, \ldots, v_{r_{\min}}$ so that for every v_q in $H_{i_q\lambda_q}$ with $1 \leq q \leq r_{\min}$ we have

$$\phi(i_q, \lambda_q, i, \lambda)(v_q) = u_q.$$

Choose middle components of $v_{r_{\min}+1}, \ldots, v_{|B|-(k-1)}$ arbitrarily. We claim that X generates T. It follows immediately from the fact that X is an $\mathcal{H}$-class transversal generating set that $E(T) \subseteq \langle X \rangle$. It now follows, by Lemma 2.28, that with this choice of $v_1, \ldots, v_{r_{\min}}$ we have $\{u_1, \ldots, u_{r_{\min}}\} \subseteq \langle X \rangle$. We conclude that $\langle \langle X \rangle \cap H_{i\lambda} \rangle = H_{i\lambda}$, which, along with the fact that X is an $\mathcal{H}$-class transversal generating set, implies by Lemma 2.4 that X generates T, and, as a consequence, generates S. Determining the size of X we have

$$|X| = |A| + |Y| = k - 1 + |B| - (k-1) = |B| = \max(|B|, r_{\min} + k - 1),$$

as required.

Case 2: $|B| < r_{\min} + k - 1$. First we note that $|B \setminus D| = |B| - (k-1) < r_{\min}$. We can extend A to $X = A \cup Y \cup Z \subseteq T$ where $A \cup Y$ has coordinates B while all $z \in Z$ have the fixed coordinate (j, μ) where $(j, \mu) \in B$. Let $Y = \{v_1, \ldots, v_\delta\} \subseteq T$ (where $\delta = |B| - (k-1)$) with coordinates $B \setminus D$ and middle components chosen so that for every v_q in $H_{i_q\lambda_q}$ we have

$$\phi(i_q, \lambda_q, i, \lambda)(v_q) = u_q.$$

Further let $Z = \{v_{\delta+1}, v_{\delta+2}, \ldots, v_{r_{\min}}\}$ with middle components chosen so that:

$$\phi(j, \mu, i, \lambda)(v_q) = u_q.$$

Then as before X generates T, and, as a consequence, generates S with

$$\begin{aligned} |X| &= |A| + |Y| + |Z| = (k-1) + \delta + (r_{\min} - (\delta + 1) + 1) \\ &= r_{\min} + k - 1 = \max(|B|, r_{\min} + k - 1) \end{aligned}$$

as required. □

Corollary 2.50. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be any completely 0-simple semigroup with connected components $I_1 \times \Lambda_1, \ldots, I_k \times \Lambda_k$. Then*

$$\operatorname{rank}(S) \leq \max(|I|, |\Lambda|, r_{\min} + k - 1).$$

Proof. Let $T = S\natural$. By Theorem 2.10 we have $\operatorname{rank}(T) = \max(|I|, |\Lambda|)$. The result now follows from Proposition 2.49. □

We now state and prove the main result of this chapter.

Theorem 2.51. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be any completely 0-simple semigroup with connected components $I_1 \times \Lambda_1, \ldots, I_k \times \Lambda_k$. Then*

$$\operatorname{rank}(S) = \max(|I|, |\Lambda|, r_{\min} + k - 1),$$

where

$$r_{\min} = \min_{\theta \in \operatorname{Map}(C^T, G)} (\operatorname{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S_{C^T,\theta}))),$$

and where $C \subseteq I \times \Lambda$ is any minimal set of component connecting coordinates and $H_{i\lambda}$ is any group $\mathcal{H}$-class.

Proof. The fact that a generating set with the required size can be found was proven in Corollary 2.50.

To complete the proof, let $A \subseteq S$ be an arbitrary generating set of S. By Lemma 2.3, we must have $|A| \geq \max(|I|, |\Lambda)$ so we are just left to show that $|A| \geq r_{\min} + k - 1$. By Lemma 2.34, the set A must have a minimal component connecting subset say $D \subseteq A$ with $|D| = k - 1$. Clearly by definition $|A \setminus D| \geq \operatorname{rank}(S : D)$. Let $T = \operatorname{Con}(S, D)$, a corresponding connected completely 0-simple semigroup, which by Corollary 2.48 satisfies

$$\operatorname{rank}(S : D) = \operatorname{rank}(T : D)$$

and $D \subseteq F(T)$. Therefore

$$\begin{aligned}
|A \setminus D| &\geq \operatorname{rank}(S : D) \\
&= \operatorname{rank}(T : D) \\
&\geq \operatorname{rank}(H_{i\lambda} : H_{i\lambda} \cap F(T)) \quad \text{(by Lemma 2.29)} \\
&\geq r_{\min}
\end{aligned}$$

and so $|A| = |D| + |A \setminus D| \geq r_{\min} + k - 1$. $\square$

2.9 Expressing $r_{\min}$ as a property of G

Recall that

$$r_{\min} = \min_{\theta \in \mathrm{Map}(C^T, G)} (\mathrm{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S_{C^T,\theta})))$$

where $S_{C^T,\theta}$ is a connected completely 0-simple semigroup. Since $S_{C^T,\theta}$ is connected we can investigate $H_{i\lambda} \cap F(S_{C^T,\theta})$ in more depth. We saw in Lemma 2.21 that if S is a connected completely 0-simple semigroup then $H_{1_I 1_\Lambda} \cap F(S)$ is isomorphic to $V_{1_I 1_\Lambda} p_{1_\Lambda 1_I}$. We will now construct a generating set for $V_{1_I 1_\Lambda} p_{1_\Lambda 1_I}$ as described in [81].

For each $i \in I$ and $\lambda \in \Lambda$ let us choose a path π_λ connecting 1_I to λ in $\Delta(P)$ and a path π_i connecting i to 1_Λ in $\Delta(P)$ (with $\pi_{1_\Lambda} = \pi_{1_I} = 1_I \to 1_\Lambda$); this is possible since S is connected. Then with

$$a_{\lambda i} = V(\pi_\lambda) p_{\lambda i} V(\pi_i) p_{1_\Lambda 1_I}$$

we have:

Lemma 2.52. *[81, Lemma 4.5] The group $V_{1_I 1_\Lambda} p_{1_\Lambda 1_I}$ is generated by the set $\mathcal{A} = \{a_{\lambda i} \mid i \in I, \lambda \in \Lambda\}$.*

Proof. Consider an arbitrary path in the graph $\Delta(P)$ starting at 1_I and ending at 1_Λ:

$$\pi = 1_I \to \lambda_1 \to i_2 \to \lambda_2 \to i_3 \to \lambda_3 \to \ldots \to i_{r-1} \to \lambda_{r-1} \to i_r \to 1_\Lambda.$$

The value of this path is:

$$p^{-1}_{\lambda_1 1_I} p_{\lambda_1 i_2} p^{-1}_{\lambda_2 i_2} p_{\lambda_2 i_3} p^{-1}_{\lambda_3 i_3} \cdots p^{-1}_{\lambda_{r-1} i_{r-1}} p_{\lambda_{r-1} i_r} p^{-1}_{1_\Lambda i_r}.$$

Now we amend the path slightly by adding in "spurs" (see Figure 2.4). These will not change the value of the path. The new path is:

$$1_I(\to 1_\Lambda \to)1_I \to \lambda_1(\xrightarrow{\pi^{-1}_{\lambda_1}} 1_I \xrightarrow{\pi_{\lambda_1}})\lambda_1 \to i_2(\xrightarrow{\pi_{i_2}} 1_\Lambda(\xrightarrow{p_{11}} 1_I \xrightarrow{p^{-1}_{11}})1_\Lambda \xrightarrow{\pi^{-1}_{i_2}})i_2 \to \lambda_2(\to \ldots$$

$$\ldots 1_I \xrightarrow{p^{-1}_{11}} 1_\Lambda \xrightarrow{\pi^{-1}_{i_r}} i_r \to 1_\Lambda.$$

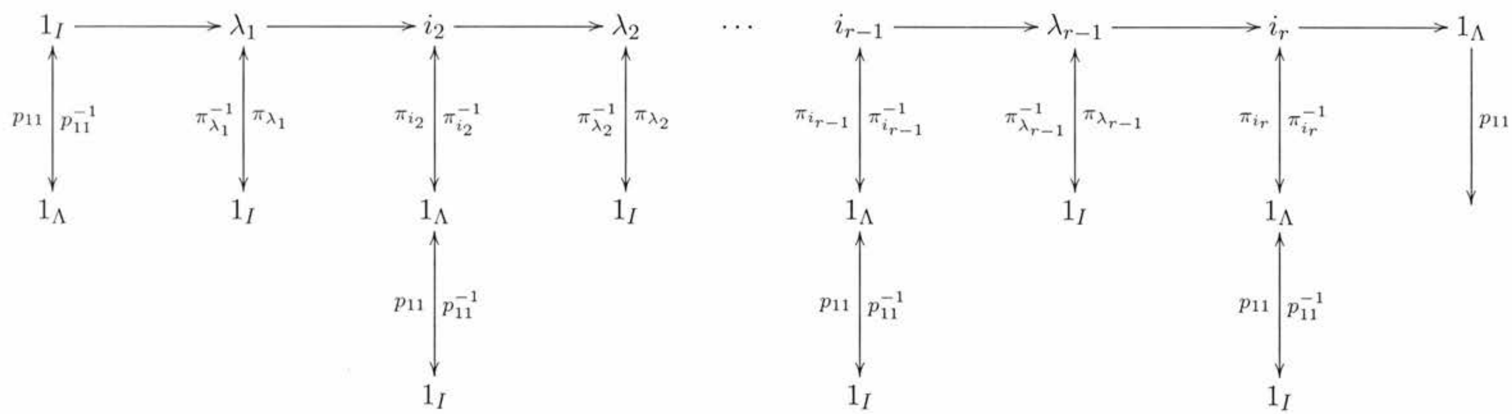


Figure 2.4: Expressing an arbitrary element of $V_{1_I 1_\Lambda}$ as a product of elements of $A = \{a_{\lambda i} : i \in I, \lambda \in \Lambda, p_{\lambda i} \neq 0\}$.

Computing the value of this path we have:

$$\begin{aligned} V(\pi)p_{1_\Lambda 1_I} &= p_{\lambda_1 1_I}^{-1} p_{\lambda_1 i_2} p_{\lambda_2 i_2}^{-1} p_{\lambda_2 i_3} p_{\lambda_3 i_3}^{-1} \cdots p_{\lambda_{r-1} i_{r-1}}^{-1} p_{\lambda_{r-1} i_r} p_{1_\Lambda i_r}^{-1} p_{1_\Lambda 1_I} \\ &= a_{\lambda_1 1_I}^{-1} a_{\lambda_1 i_2} a_{\lambda_2 i_2}^{-1} a_{\lambda_2 i_3} a_{\lambda_3 i_3}^{-1} \cdots a_{\lambda_{r-1} i_{r-1}}^{-1} a_{\lambda_{r-1} i_r} a_{1_\Lambda i_r}^{-1}. \end{aligned}$$

Therefore the group $V_{1_I 1_\Lambda} p_{1_\Lambda 1_I}$ is generated by $\mathcal{A}$. □

Using this generating set we may obtain a more concrete result than Theorem 2.51. It generalises [81, Theorem 4.6].

Theorem 2.53. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a finite Rees matrix semigroup with k connected components $I_1 \times \Lambda_1, \ldots, I_k \times \Lambda_k$. For every $j = 1, \ldots, k$ choose $(1_{I_j}, 1_{\Lambda_j}) \in I_j \times \Lambda_j$ with $p_{1_{\Lambda_j} 1_{I_j}} \neq 0$ so that $H_{1_{I_j}, 1_{\Lambda_j}}$ are group $\mathcal{H}$-classes. For $\lambda \in \Lambda_r$ and $i \in I_l$ let π_λ be a path connecting 1_{I_r} to λ in the subgraph $I_r \cup \Lambda_r$ and let π_i be a path connecting i to 1_{Λ_l} in the subgraph $I_l \cup \Lambda_l$ (with $\pi_{1_{I_q}} = \pi_{1_{\Lambda_q}} = 1_{I_q} \xrightarrow{p_{1_{\Lambda_q} 1_{I_q}}{}^{-1}} 1_{\Lambda_q}$). For every $r = 1, \ldots, k$ let*

$$a_{\lambda i} = V(\pi_\lambda) p_{\lambda i} V(\pi_i) p_{1_{\Lambda_r} 1_{I_r}} \quad ((\lambda, i) \in \Lambda_r \times I_r).$$

Let H_r be the subgroup of G generated by the set $\{a_{\lambda i} \mid (\lambda, i) \in \Lambda_r \times I_r, a_{\lambda i} \neq 0\}$. Then

$$\operatorname{rank}(S) = \max(|I|, |\Lambda|, \rho_{\min} + k - 1),$$

where

$$\rho_{\min} = \min\{\operatorname{rank}(G : \bigcup_{i=1}^{k} g_i H_i g_i{}^{-1}) | g_1, \ldots, g_k \in G\}.$$

Proof. The way to think about this complicated looking theorem is as follows. With each connected component $\Lambda_i \times I_i$ of P there is an associated subgroup H_i of G which is determined by the position and values of the entries of $\Lambda_i \times I_i$. Note that as a consequence of Lemmas 2.21 and 2.52 the subgroup H_r is isomorphic to $F(S) \cap H_{i\lambda}$ where $H_{i\lambda}$ is some fixed group $\mathcal{H}$-class of component $\Lambda_r \times I_r$. We conjugate each of these subgroups in turn by any elements we want from G. We go on to look at the relative rank in G of the union of these subgroups. The number $\rho_{\min}$ is the smallest possible value this relative rank can take.

Let C be the connecting set

$$C = \{(1_{I_2}, 1_{\Lambda_1}), (1_{I_3}, 1_{\Lambda_1}), (1_{I_4}, 1_{\Lambda_1}), \ldots (1_{I_k}, 1_{\Lambda_1})\} \subseteq I \times \Lambda,$$

let $\theta \in \operatorname{Map}(C^T, G)$ with $(1_{\Lambda_1}, 1_{I_r})\theta = g_r$ for $2 \leq r \leq k$ with $g_1 = p_{1_{\Lambda_1} 1_{I_1}}$ and

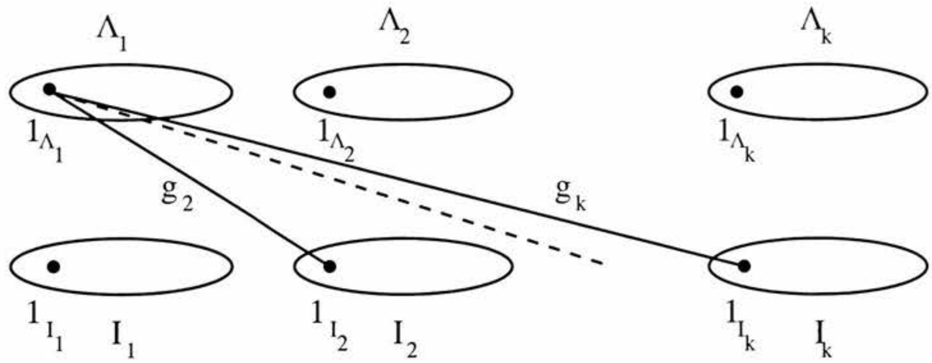


Figure 2.5: The connected graph $\Delta(Q)$ constructed from the disconnected graph $\Delta(P)$ and the connecting set C.

consider

$$\operatorname{rank}(H_{1_{I_1}1_{\Lambda_1}} : H_{1_{I_1}1_{\Lambda_1}} \cap F(S_{C^T,\theta})).$$

We are really interested in $H_{1_{I_1}1_{\Lambda_1}} \cap F(T)$, where $T = S_{C^T,\theta}$ is connected and so we can use Lemmas 2.21 and 2.52 to compute it.

In order to determine $H_{1_{I_1}1_{\Lambda_1}} \cap F(T)$ we consider the graph $\Delta(Q)$ (see Figure 2.5) where $Q = P_{C^T,\theta}$ with entries $q_{\lambda i}, i \in I, \lambda \in \Lambda$. The graph $\Delta(Q)$ is connected and in it we can find paths from 1_{I_1} to any λ and from any i to 1_{Λ_1} which we need in order to compute $H_{1_{I_1}1_{\lambda_1}} \cap F(T)$. For each $\lambda \in \Lambda_r \subseteq \Lambda$ let ϖ_λ be the following path from 1_{I_1} to λ

$$\varpi_\lambda : 1_{I_1} \to 1_{\Lambda_1} \to 1_{I_r} \xrightarrow{\pi_r} \lambda.$$

Note that $V(\varpi_\lambda) = q^{-1}_{1_{\Lambda_1}1_{I_1}} g_r V(\pi_r)$. Similarly for $i \in I_r \subseteq I$ let ϖ_i be the following path from i to 1_{Λ_1}

$$\varpi_i : i \xrightarrow{\pi_i} 1_{\Lambda_r} \to 1_{I_r} \to 1_{\Lambda_1}.$$

Note that $V(\varpi_i) = V(\pi_i) q_{1_{\Lambda_r}1_{I_r}} {g_r}^{-1}$. Also choose $\varpi_{1_{I_1}} = \varpi_{1_{\Lambda_1}} = 1_{I_1} \to 1_{\Lambda_1}$. Now let

$$w_{\lambda i} = V(\varpi_\lambda) q_{\lambda i} V(\varpi_i) q_{1_{\Lambda_1}1_{I_1}}.$$

For $(\lambda, i) \in \Lambda_m \times I_n$ with $m \neq n$ we either have $q_{\lambda i} = 0$ in which case $w_{\lambda i} = 0$ or else we have $\lambda = 1_{\Lambda_1}, i = 1_{I_n}$ and $q_{\lambda i} = q_{1_{\Lambda_1}1_{I_n}} = g_n$ in which case

$$w_{\lambda i} = w_{1_{\Lambda_1}1_{I_r}} = V(\varpi_{1_{\Lambda_1}}) g_r V(\varpi_{1_{I_r}}) q_{1_{\Lambda_1}1_{I_1}} = {q_{1_{\Lambda_1}1_{I_1}}}^{-1} g_r {g_r}^{-1} q_{1_{\Lambda_1}1_{I_1}} = 1_G.$$

When $(\lambda, i) \in \Lambda_r \times I_r$ for some $1 \leq r \leq k$, we have

$$\begin{aligned} w_{\lambda i} &= V(\varpi_\lambda) q_{\lambda i} V(\varpi_i) q_{1_{\Lambda_1} 1_{I_1}} \\ &= (q_{1_{\Lambda_1} 1_{I_1}}{}^{-1} g_r) V(\pi_r) q_{\lambda i} V(\pi_i) q_{1_{\Lambda_r} 1_{I_r}} (g_r{}^{-1} q_{1_{\Lambda_1} 1_{I_1}}) \\ &= (p_{1_{\Lambda_1} 1_{I_1}}{}^{-1} g_r) V(\pi_r) p_{\lambda i} V(\pi_i) p_{1_{\Lambda_r} 1_{I_r}} (g_r{}^{-1} p_{1_{\Lambda_1} 1_{I_1}}) \\ &= h_r a_{\lambda i} h_r{}^{-1} \end{aligned}$$

with $h_r = p_{1_{\Lambda_1} 1_{I_1}}{}^{-1} g_r$ and $a_{\lambda i}$ defined as before with $(\lambda, i) \in \Lambda_r \times I_r$.

It follows that the subgroup generated by

$$\{w_{\lambda i} \mid \lambda \in \Lambda, i \in I, w_{\lambda i} \neq 0\}$$

is equal to the subgroup generated by

$$\bigcup_{j=1}^{k} \{h_j a_{\lambda i} h_j{}^{-1} | (\lambda, i) \in \Lambda_j \times I_j\}.$$

Therefore

$$\begin{aligned} r_{\min} &= \min_{\theta \in \mathrm{Map}(C^T, G)} (\mathrm{rank}(H_{1_{I_1} 1_{\Lambda_1}} : H_{1_{I_1} 1_{\Lambda_1}} \cap F(S_{C^T, \theta}))) \\ &= \min_{\theta \in \mathrm{Map}(C^T, G)} (\mathrm{rank}(G : V_{1_{I_1} 1_{\Lambda_1}} p_{1_{\Lambda_1} 1_{I_1}})) \\ &= \min_{\theta \in \mathrm{Map}(C^T, G)} (\mathrm{rank}(G : \langle w_{\lambda i} \mid \lambda \in \Lambda, i \in I, w_{\lambda i} \neq 0 \rangle)) \\ &= \min\{\mathrm{rank}(G : \bigcup_{i=1}^{k} g_i H_i g_i{}^{-1}) \mid g_2, \ldots, g_k \in G\} \quad \text{with } g_1 = p_{1_{\Lambda_1} 1_{I_1}} \\ &= \min\{\mathrm{rank}(G : \bigcup_{i=1}^{k} g_i H_i g_i{}^{-1}) \mid g_1, \ldots, g_k \in G\} \\ &= \rho_{\min}. \end{aligned}$$

□

2.10 Isomorphism theorem and normal forms

Given a finite 0-simple semigroup it will have, up to isomorphism, a unique non-zero maximal subgroup. This subgroup is isomorphic to G in the Rees matrix construction. The $\mathcal{R}$-classes are indexed by I and the $\mathcal{L}$-classes by Λ. The matrix P with entries in $G \cup \{0\}$ is not, however, uniquely determined by S. In the proof of the Rees theorem (see Chapter 1) there is some choice in its construction.

That is, with G, I and Λ fixed, two different matrices may give rise to isomorphic semigroups. The following result demonstrates exactly how much variation in the matrix P is allowed.

Theorem 2.54. *[57, Theorem 3.4.1] Two regular Rees matrix semigroups $S = \mathcal{M}^0[G; I, \Lambda; P]$ and $T = \mathcal{M}^0[K; J, M; Q]$ are isomorphic if and only if there exist an isomorphism $\theta : G \to K$, bijections $\psi : I \to J$, $\chi : \Lambda \to M$ and elements u_i $(i \in I), \nu_\lambda$ $(\lambda \in \Lambda)$ such that:*

$$p_{\lambda i}\theta = \nu_\lambda q_{\lambda\chi, i\psi} u_i$$

for all $i \in I$ and $\lambda \in \Lambda$.

It follows from Theorem 2.54 that if $S = \mathcal{M}^0[G; I, \Lambda; P]$ and $T = \mathcal{M}^0[K; J, M; Q]$ are isomorphic then so are the graphs $\Delta(P)$ and $\Delta(Q)$. The values of the paths in $\Delta(P)$ will however not equal the values of the corresponding paths in $\Delta(Q)$, in general.

Another way of expressing this result is the following. If there is a $\Lambda \times \Lambda$ diagonal matrix V with entries in G and an $I \times I$ diagonal matrix with entries in G such that $P = VQU$, then the map $(i, a, \lambda) \mapsto (i, u_i a v_\lambda, \lambda)$ is an isomorphism from $\mathcal{M}^0[G; I, \Lambda; P]$ onto $\mathcal{M}^0[G; I, \Lambda; Q]$.

In particular, when P happens to contain only non-zero entries, i.e. $S = J \cup \{0\}$ where J is a completely simple semigroup, then there is an obvious normal form that any matrix P may be put into:

Definition 2.55. We say that the matrix $P = (p_{\lambda i})$ with entries in G is in *normal* form if every entry in the first row and the first column is equal to the identity 1_G of G.

Theorem 2.56. *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.*

If S is a connected completely 0-simple semigroup then we can normalise the matrix in a special way using the graph $\Delta(P)$.

Theorem 2.57. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a connected completely 0-simple semigroup. For any spanning tree $\mathcal{T}$ of the bipartite graph $\Delta(P)$ it is possible to normalize the matrix P so that every edge in $\mathcal{T}$ is labelled by the identity of G.*

Proof. We proceed by induction on the number of vertices of the graph $\Delta(P)$. When $\Delta(P)$ has two vertices the result is trivial. For the inductive step we will isolate a leaf in the tree, remove it, normalize the resulting completely 0-simple semigroup, then add the vertex back in and normalize appropriately.

Let $\mathcal{T}$ be a spanning tree of the graph $\Delta(P)$. Let $v \in \mathcal{T}$ be a vertex of degree one in $\mathcal{T}$ (a leaf of the tree) and suppose that $v \in I$. The vertex v labels a column of the matrix P and by construction the matrix $P' = P[0][v]$ (the matrix P with column v removed) is also connected. By induction the matrix P' may be normalized in such a way that the edges of $\mathcal{T} \setminus \{v\}$, which is a spanning tree of $\Delta(P')$, all have value 1_G. Let $\theta : \Lambda \to G$ and $\delta : I \setminus \{v\} \to G$ be the functions defining such a normalization. We will extend δ to δ^* so that the pair (θ, δ^*) normalize P in the desired way. Since v has degree one in $\mathcal{T}$ there is a unique element $\lambda_v \in \Lambda$ to which i is connected in $\mathcal{T}$. Therefore $p_{\lambda_v,v} \neq 0$ in P. Now define

$$\delta^*(v) = p^{-1}_{\lambda_v,v}\theta(\lambda_v)^{-1},$$

and thus the edge (λ_v, v) has value:

$$(\theta(\lambda_v))p_{\lambda_v,v}(p^{-1}_{\lambda_v,v}\theta(\lambda_v)^{-1}) = 1_G.$$

Also, since v has degree one in $\mathcal{T}$, all of the other values of edges of $\mathcal{T}$ in $\Delta(P)$ once normalized by (θ, δ^*) are left unaltered (i.e. are all equal to 1_G). This completes the inductive step. There is a dual argument when $v \in \Lambda$. □

Now consider the generating set $\mathcal{A} = \{a_{\lambda i} \mid i \in I, \lambda \in \Lambda\}$ described in the previous section. It is constructed by fixing paths π_i for each $i \in I$ and π_λ for each $\lambda \in \Lambda$. Once P has been normalized, by the process described above, all of the paths π_i and π_λ can be chosen to only use edges from the spanning tree $\mathcal{T}$. Therefore $V(\pi_\lambda) = V(\pi_i) = 1_G$ for all i and λ. In addition to this we can find H_{11} such that $p_{11} = 1_G$. Now

$$\begin{aligned} \mathcal{A} &= \{a_{\lambda i} \mid i \in I, \lambda \in \Lambda\} = \{V(\pi_\lambda)p_{\lambda i}V(\pi_i)p_{11} \mid i \in I, \lambda \in \Lambda\} \\ &= \{p_{\lambda i} \mid i \in I, \lambda \in \Lambda\}. \end{aligned}$$

We conclude that the group $V_{11}p_{11}$ is equal to $\langle p_{\lambda i} \mid i \in I, \lambda \in \Lambda, p_{\lambda i} \neq 0\rangle$, the subgroup generated by the non-zero entries in the matrix P. Carrying out this normalization process on each of the connected components of a matrix P in turn gives a matrix that we say is in *Graham normal form*. Graham [41] was the first to realise that the matrix P may be normalized in such a way.

Theorem 2.58. *[41, Theorem 2] Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a finite regular Rees matrix semigroup. It is always possible to normalize the structure matrix P to* *obtain Q with the following properties:*

(i) *the matrix Q is a direct sum of r blocks $C_1, \ldots, C_r$ suggested by the following picture:*

$$\begin{array}{c} A_1 \\ A_2 \\ \vdots \\ A_r \end{array} \overset{\begin{array}{cccc} B_1 & B_2 & \ldots & B_r \end{array}}{\begin{pmatrix} C_1 & & & 0 \\ & C_2 & & \\ & & \ddots & \\ 0 & & & C_r \end{pmatrix}}.$$

(ii) *Each matrix $C_i : A_i \times B_i \to G^0$ is regular and*

$$F(S) = \langle E(S) \rangle = \bigcup_{i=1}^{r} \mathcal{M}^0[G_i; A_i, B_i; C_i]$$

where G_i is the subgroup of G generated by the non-zero entries of C_i, for $i = 1, \ldots, r$.

2.11 Applications, examples and remarks

We begin by using the notion of Graham normal form to obtain the following neat formulation of Theorem 2.53.

Corollary 2.59. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a finite Rees matrix semigroup with k connected components $I_1 \times \Lambda_1, \ldots, I_k \times \Lambda_k$ and with regular matrix P in Graham normal form. For every $r = 1, \ldots, k$ let H_r be the subgroup of G generated by the non-zero entries of component $C_r = I_r \times \Lambda_r$ of the matrix P. Then*

$$\text{rank}(S) = \max(|I|, |\Lambda|, \sigma_{\min} + k - 1)$$

where

$$\sigma_{\min} = \min\{\text{rank}(G : \bigcup_{i=1}^{k} g_i H_i {g_i}^{-1}) \mid g_1, \ldots, g_k \in G\}.$$

Proof. By Theorem 2.57 we may choose, for every $j = 1, \ldots, k$, an element $(1_{I_j}, 1_{\Lambda_j}) \in I_j \times \Lambda_j$ with $p_{1_{\Lambda_j}, 1_{I_j}} = 1_G$. For every $r = 1, \ldots, k$ we have

$$a_{\lambda i} = V(\pi_\lambda) p_{\lambda i} V(\pi_i) 1_G.$$

The subgroup H_r of G is generated by:

$$\{a_{\lambda i} \mid (\lambda, i) \in \Lambda_r \times I_r, p_{\lambda i} \neq 0\}.$$

In fact, by Lemma 2.52

$$H_r = V_{1_{I_r}1_{\Lambda_r}}1_G = V_{1_{I_r}1_{\Lambda_r}}.$$

By Lemma 2.21 the map $\psi : H_{1_{I_r}1_{\Lambda_r}} \to G$ defined by

$$\psi(1_{I_r}, g, 1_{\Lambda_r}) = g$$

is an isomorphism which maps $H_{1_{I_r}1_{\Lambda_r}} \cap F(S)$ onto $V_{1_{I_r}1_{\Lambda_r}}$. Since the matrix P is in Graham normal form we have

$$H_{1_{I_r}1_{\Lambda_r}} \cap F(S) = 1_{I_r} \times G_r \times 1_{\Lambda_r}$$

where G_r is the subgroup of G generated by the non-zero entries in component C_r. We conclude that

$$H_r = V_{1_{I_r}1_{\Lambda_r}} = G_r \subseteq G.$$

In other words, H_r and G_r are equal as sets. Now the result follows as a direct application of Theorem 2.53. □

In fact, using Theorem 2.54, once the best conjugating elements have been found we can conjugate the entries in the matrix putting it in a new normal form.

Corollary 2.60. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a finite completely 0-simple semigroup where P has k connected components. It is possible to normalize the matrix P to obtain a matrix Q with the property that:*

$$\text{rank}(S) = \max(|I|, |\Lambda|, \text{rank}(G : K) + k - 1)$$

where K is the subgroup of G generated by the non-zero entries in the matrix Q.

Note that the above result contains less information than the main theorems, Theorems 2.51 and 2.53, in the sense that there are no instructions given for finding the appropriate normalization. It just tells us that such a normalization exists.

A *Hamiltonian* group is a group all of whose subgroups are normal. In particular, all abelian groups are Hamiltonian. The quaternion group Q_8 is an example of a non-abelian Hamiltonian group. If G is an abelian group, or more generally a Hamiltonian group, then conjugating has no effect on subgroups which gives us the following corollary.

Corollary 2.61. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a finite Rees matrix semigroup with G a Hamiltonian group, k connected components and with regular matrix P in Graham normal form. Let H be the subgroup of G generated by the non-zero entries of P. Then*

$$\operatorname{rank}(S) = \max(|I|, |\Lambda|, \operatorname{rank}(G : H) + k - 1).$$

Given a completely 0-simple semigroup it is sometimes possible to express it as a Rees matrix semigroup over a group where P is made up entirely of elements from $\{0, 1_G\}$. A formula for the rank, in this situation, is given by the following corollary.

Corollary 2.62. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a finite Rees matrix semigroup with k connected components and with regular matrix P only containing entries from $\{0, 1_G\}$. Then*

$$\operatorname{rank}(S) = \max(|I|, |\Lambda|, \operatorname{rank}(G) + k - 1).$$

Proof. Since P has only 0 and 1_G as entries then it is in Graham normal form. Therefore

$$r_{\min} = \min\{\operatorname{rank}(G : \bigcup_{i=1}^{k} g_i H_i {g_i}^{-1}) | g_1, \ldots, g_k \in G\} = \operatorname{rank}(G : \{1_G\}) = \operatorname{rank}(G).$$

□

Recall that the Brandt semigroup $B = B(G, \{1, \ldots, n\})$ is a Rees matrix semigroup $\mathcal{M}^0[G; I, I; P]$ where $P \sim I_n$, the $n \times n$ identity matrix, and $I = \{1, \ldots, n\}$. The Brandt semigroup represents the extreme case where $\Gamma(\mathbb{H}_S)$ is as disconnected as possible. The Brandt semigroup clearly satisfies the criteria of Corollary 2.62 and as a result we obtain the following theorem of Gomes and Howie:

Corollary 2.63. *[37, Theorem 3.3] Let $B = B(G, \{1, \ldots, n\})$ be a Brandt semigroup, where G is a finite group of rank $r \geq 1$. Then the rank of B is $r + n - 1$.*

We saw in Lemma 2.9 that when $\operatorname{rank}(G) = 0$ (i.e. G is the trivial group) then $\operatorname{rank}(B) = n$. Note that the original result of Gomes and Howie actually gave the rank of B as an inverse semigroup but in [35] Garba shows that, when $r \geq 1$, the rank of B as a semigroup is equal to the rank of B as an inverse semigroup.

Definition 2.64. We call S *locally generated* if $H_i = G$ for some $1 \leq i \leq k$.

Corollary 2.65. *If $S = \mathcal{M}^0[G; \Lambda, I; P]$ is a locally generated completely 0-simple semigroup then*

$$\mathrm{rank}(S) = \max(|I|, |\Lambda|).$$

Proof.

$$\begin{aligned}\mathrm{rank}(S) &= \max(|I|, |\Lambda|, \min\{\mathrm{rank}(G : \bigcup_{i=1}^{k} g_i H_i {g_i}^{-1}) | g_1, \ldots, g_k \in G\})) \\ &= \max(|I|, |\Lambda|, \mathrm{rank}(G : G)) \\ &= \max(|I|, |\Lambda|).\end{aligned}$$

□

After proving Theorem 2.10, and considering the results for the connected case, it would perhaps have seemed reasonable, as an initial guess, to think that the rank of a completely 0-simple semigroup might be given by

$$\mathrm{rank}(S) = \max(|I|, |\Lambda|, \min_{i \in I, \lambda \in \Lambda}(\mathrm{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S)))). \quad (2.1)$$

In fact the expression on the right hand side does not even serve as an upper or lower bound for the rank in general, as the following two examples show.

Example 2.66. Let G be a group with $\mathrm{rank}(G) = 3$ (such as $C_2 \times C_2 \times C_2$ for example) and let S be the Brandt semigroup $S = B(G, \{1, 2\})$. Then $F(S) = \{(1, 1_G, 1), (2, 1_G, 2), 0\}$ and so

$$\begin{aligned}\max(|I|, |\Lambda|, \min_{i \in I, \lambda \in \Lambda}(\mathrm{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S)))) &= \max(2, 2, 1) = 2 < 4 \\ &= r + k - 1 = \mathrm{rank}(S).\end{aligned}$$

Example 2.67. Let G be a group with $\mathrm{rank}(G) = 11$ (such as ${C_2}^{11}$ for example) and suppose $G = \langle g_1, \ldots, g_{11} \rangle$. Let

$$P = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & g_1 & g_2 & 0 & 0 & 0 \\ 1 & g_3 & g_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & g_5 & g_6 \\ 0 & 0 & 0 & 1 & g_7 & g_8 \end{pmatrix}$$

which is in fact in Graham normal form. Letting $H_1 = \langle g_1, g_2, g_3, g_4 \rangle$ and $H_2 =$

$\langle g_5, g_6, g_7, g_8 \rangle$ we have

$$r_{\min} = \min\{\text{rank}(G : h_1 H_1 h_1^{-1} \cup h_2 H_2 h_2^{-1}) | h_1, h_2 \in G\} \leq \text{rank}(G : H_1 \cup H_2) = 3.$$

Then we have

$$\text{rank}(S) = \max(|I|, |\Lambda|, r_{\min} + 2 - 1) = \max(6, 6, r_{\min} + 1) = 6.$$

The right hand side of (2.1) gives

$$\begin{aligned} \max(|I|, |\Lambda|, \min_{i\in I, \lambda \in \Lambda}(\text{rank}(H_{i\lambda} : H_{i\lambda} \cap F(S))) &= \max(6, 6, \min(7, 7)) \\ &= 7 > 6 = \text{rank}(S). \end{aligned}$$

2.12 Non-regular Rees matrix semigroups

The definition of Rees matrix semigroup generalizes in a number of ways. For example, if we remove the assumption that the matrix P is regular we still obtain a semigroup, but in general it will not be a 0-simple semigroup. We call such semigroups *generalized Rees matrix semigroups*. We begin this section with an easy generalization of the main theorem of the previous section with the removal of the assumption that the matrix is regular.

Theorem 2.68. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a generalized Rees matrix semigroup over a group G. Write P in the form:*

$$\begin{array}{cc} & \begin{array}{cc} I_1 & I_2 \end{array} \\ \begin{array}{c} \Lambda_1 \\ \\ \Lambda_2 \end{array} & \left(\begin{array}{c|c} Q & \\ \hline & \end{array}\right) \end{array}$$

where $Q = \Lambda_1 \times I_1$ is regular and the rest of the entries equal zero. Let $T = \mathcal{M}^0[G; I_1, \Lambda_1; Q]$, which is a completely 0-simple semigroup. Then

$$\text{rank}(S) = \text{rank}(T) + |I_2| + |\Lambda_2|.$$

Proof. ($\leq$) We begin by demonstrating that a generating set with the required size can be found. Let $B \subseteq I_1 \times G \times \Lambda_1$ be a generating set for T with $|B| = \text{rank}(T)$.

Fix $i^* \in I_1$ and $\lambda^* \in \Lambda_1$ and define

$$B_I = \{(i^*, 1_G, \lambda) : \lambda \in \Lambda_2\}, \quad B_\Lambda = \{(i, 1_G, \lambda^*) : i \in I_2\}.$$

Now

$$|B \cup B_I \cup B_\Lambda| = \text{rank}(T) + |I_2| + |\Lambda_2|$$

and this set generates S.

($\geq$) Let A be a generating set of S with $|A| = \text{rank}(S)$.

Claim. $\langle A \cap T \rangle = \langle A \rangle \cap T = T$.

Proof. Consider the graph $\Delta(S : A)$ and recall the connection between $\langle A \rangle$ and the values of the valid paths in the graph. Consider a path (with non-zero value) beginning in I_1 and ending in Λ_1. No such path passes through a vertex of $I_2 \cup \Lambda_2$, since there are no directed edges from Λ_2 to I_2, from Λ_1 to I_2, or from Λ_2 to I_1. □

Claim. $\langle A \setminus (I_2 \times G \times \Lambda_2) \rangle = S$.

Proof. Let $B = A \setminus (I_2 \times G \times \Lambda_2)$. First observe that for every $s \in I_2 \times G \times \Lambda_2$ and for every $t \in S$ we have $st = ts = 0$. It follows from this that since A generates S we must have $\langle B \rangle \supseteq S \setminus (I_2 \times G \times \Lambda_2)$. Finally note that $S \setminus (I_2 \times G \times \Lambda_2)$ generates S since for $i_2 \in I_2$ and $\lambda_2 \in \Lambda_2$, if we fix $\lambda \in \Lambda_1$ and $i \in I_1$ such that $p_{\lambda i} \neq 0$, which we can since Q is regular, we have:

$$\{(i_2, g, \lambda_2) : g \in G\} = \{(i_2, g, \lambda) : g \in G\}\{(i, g, \lambda_2) : g \in G\}.$$

□

Returning to the proof of the theorem, since $A \setminus (I_2 \times G \times \Lambda_2)$ generates S it must intersect every row and every column of S. (Notice here that it is incorrect to use the expressions $\mathcal{R}$- and $\mathcal{L}$-class for the rows and the columns since these are not the $\mathcal{R}$-classes and $\mathcal{L}$-classes anymore.) We conclude that:

$$\begin{aligned} \text{rank}(S) &= |A| \geq |A \cap T| + |A \cap (I_2 \times G \times \Lambda_1)| + |A \cap (I_1 \times G \times \Lambda_2)| \\ &\geq \text{rank}(T) + |I_2| + |\Lambda_2| \end{aligned}$$

as required. □

In particular, if $S = \mathcal{M}^0[G; I, \Lambda; P]$ is a finite completely 0-simple semigroup with $I' \subseteq I$ and $\Lambda' \subseteq \Lambda$, then $(I' \times G \times \Lambda') \cup \{0\}$ is a subsemigroup of S and is isomorphic to the generalized 0-Rees matrix semigroup $\mathcal{M}^0[G; I', \Lambda'; Q]$ where $q_{\lambda i} = p_{\lambda i}$ for all $(\lambda, i) \in \Lambda' \times I'$. In other words, if S is a finite completely 0-simple semigroup and if $R_1, \ldots, R_m$ is a set of non-zero $\mathcal{R}$-classes of S, and $L_1, \ldots, L_n$ is a set of non-zero $\mathcal{L}$-classes of S then

$$\{(R_1 \cup \ldots \cup R_m) \cap (L_1 \cup \ldots \cup L_n)\} \cup \{0\}$$

is a subsemigroup of S and is isomorphic to a generalized 0-Rees matrix semigroup, as described above.

2.13 A family of transformation semigroups

As mentioned in the introduction, in current literature one of the most common areas of investigation has been that of the ranks of various semigroups of transformations. Many examples that have been considered happen to be idempotent generated and therefore the question of rank is answered simply by applying Theorem 2.13. In order to find less trivial applications for Theorem 2.51 examples that are not idempotent generated, and those that are not even connected, should be considered.

Definition 2.69. Let $n, r \in \mathbb{N}$ with $2 < r < n$. Let A be a set of r-subsets of $\{1, \ldots, n\}$ and let B be a set of partitions of $\{1, \ldots, n\}$, each with r classes. Define:

$$S(A, B) = \langle\ \{\alpha \in T_n : \operatorname{im} \alpha \in A, \ker \alpha \in B\}\ \rangle$$

the semigroup generated by all maps with image in A and kernel in B.

Example 2.70. In [68], [67] and [66] subsemigroups of T_n generated by elements all with the same kernel type are considered. These semigroups are idempotent generated and in this series of papers their ranks and idempotent ranks are computed. The semigroups $S(A, B)$ are a more general class since when $A = \{X \subseteq \{1, \ldots, n\} : |X| = r\}$ and B is the set of kernels of a particular partition type then we recover the examples of [68].

Example 2.71. Let S be a transformation semigroup over a finite set X generated by elements with fixed rank r and with the property that S has a subgroup that is isomorphic to the symmetric group S_r. Then this semigroup belongs to the class of semigroups of the type $S(A, B)$.

Definition 2.72. Let Γ be a finite simple graph. For a subset X of the vertices of Γ define

$$V_0(X) = \{x \in X : d(x) = 0\},$$

the set of all isolated vertices, and

$$V_+(X) = \{x \in X : d(x) > 0\},$$

the vertices with non-zero degree so that $X = V_0(X) \cup V_+(X)$. Also define $v_0(X) = |V_0(X)|$, $v_+(X) = |V_+(X)|$ and $\mathcal{MD}(\Gamma) = \max\{d(v) : v \in V(\Gamma)\}$: the maximum degree of a vertex of the graph.

Recall that the edges in the graph $\Delta(P)$ corresponded to idempotents in the completely 0-simple semigroup. From Proposition 2.14 the idempotents in T_n are indexed by pairs (I, K) where I is an image and K is a kernel such that I is a transversal of K. This leads to the following definition.

Definition 2.73. Let A be a set of r-subsets of $\{1, \ldots, n\}$ and B be a set of partitions of $\{1, \ldots, n\}$ of weight r. Define the bipartite graph $\Gamma(A, B)$ to have vertices $A \cup B$ and $a \in A$ connected to $b \in B$ if and only if a is a transversal of b.

Lemma 2.74. *Let $\mathcal{I} = \{\alpha \in S(A, B) : |\operatorname{im} \alpha| < r\}$ which, provided it is non-empty, is a two-sided ideal of $S(A, B)$. Then $S(A, B)/\mathcal{I}$ is isomorphic to a generalized Rees matrix semigroup over the symmetric group $G \cong S_r$ and*

$$\operatorname{rank}(S(A, B)) = \operatorname{rank}(S(A, B)/\mathcal{I}).$$

Proof. Let $S = K(n, r)$. Let $T \cong K(n, r)/K(n, r-1)$ which is a finite completely 0-simple semigroup. Let $\phi : S \to T$ be the natural epimorphism from S onto T. By definition

$$S(A, B)/\mathcal{I} = (S(A, B) \cap D_r) \cup \{0\}$$

with a product of two elements $x, y \in S(A, B) \cap D_r$ equal to 0 if and only if $xy \notin S(A, B) \cap D_r$. By the definition of ϕ it follows that $S(A, B)/\mathcal{I}$ is isomorphic to $(S(A, B) \cap D_r)\phi \cup \{0\} \leq T$. From the definition of the semigroup $S(A, B)$ this subsemigroup of T has the form

$$\{(R_1 \cup \ldots \cup R_m) \cap (L_1 \cup \ldots \cup L_n)\} \cup \{0\}$$

where $R_1, \ldots, R_m$ is a set of non-zero $\mathcal{R}$-classes of S and $L_1, \ldots, L_n$ is a set of non-zero $\mathcal{L}$-classes of S. Therefore, if $T \cong \mathcal{M}^0[G; I, \Lambda; P]$ then there are subsets

$I' \subseteq I$ and $\Lambda' \subseteq \Lambda$ satisfying

$$(S(A,B) \cap D_r)\phi \cup \{0\} \cong (I' \times G \times \Lambda') \cup \{0\}$$

which is a generalized 0-Rees matrix semigroup. Since the maximal subgroups of D_r are isomorphic to S_r it follows that $G \cong S_r$.

For the second part, by definition $S(A,B) \cap D_r$ generates $S(A,B)$. Therefore $\text{rank}(S(A,B)) \leq \text{rank}((S(A,B)/\mathcal{I}))$. On the other hand, since $S(A,B)/\mathcal{I}$ is a homomorphic image of $S(A,B)$ it follows that $\text{rank}(S(A,B)) \geq \text{rank}((S(A,B)/\mathcal{I}))$. □

Theorem 2.75. *Let $n, r \in \mathbb{N}$ with $2 < r < n$. Let A be a set of r-sets of $\{1, \dots, n\}$ and B be a set partitions of $\{1, \dots, n\}$ with weight r, and construct the bipartite graph $\Gamma(A,B)$. Then:*

$$\text{rank}(S(A,B)) = \begin{cases} \max(v_+(A), v_+(B)) + v_0(A \cup B) & \text{if } \mathcal{MD}(A \cup B) \geq 2 \\ \max(v_+(A), v_+(B)) + v_0(A \cup B) + 1 & \text{if } \mathcal{MD}(A \cup B) = 1 \\ |A||B|r! & \text{if } \mathcal{MD}(A \cup B) = 0 \end{cases}$$

where v_+, v_0 and $\mathcal{MD}(A \cup B)$ refer to values of the graph $\Gamma(A,B)$.

Proof. As a result of Lemma 2.74 is is sufficient to prove the result for the generalized Rees matrix semigroup $S = \mathcal{M}^0[S_r; I, \Lambda; P]$ which is isomorphic to $S(A,B)/\mathcal{I}$. There are three cases to consider depending on the value of $\mathcal{MD}(A \cup B)$.

Case 1: $\mathcal{MD}(A \cup B) = 0$. In this case the structure matrix P consists entirely of zeros so $S(A,B)/\mathcal{I}$ is a null semigroup with $|A||B|r!$ non-zero elements and the result follows trivially.

Case 2: $\mathcal{MD}(A \cup B) = 1$. In this case the matrix P has the form:

$$\begin{array}{c} \\ \Lambda_1 \\ \\ \Lambda_2 \end{array} \begin{array}{c} \begin{array}{cc} I_1 & I_2 \end{array} \\ \left(\begin{array}{c|c} D & \\ \hline & \end{array} \right) \end{array}$$

where D is a diagonal matrix with entries in G. As in Theorem 2.68 let $T = \mathcal{M}^0[G; I_1, \Lambda_1; D]$ which, since D is diagonal, is isomorphic to the Brandt semigroup $B(G, |I_1|)$ with $G \cong S_r$. Since $r \geq 3$ it follows that $\text{rank}(G) = \text{rank}(S_r) = 2$. Thus, by Corollary 2.63 we have:

$$\text{rank}(T) = |I_1| + \text{rank}(G) - 1 = |I_1| + 2 - 1 = |I_1| + 1.$$

Now by definition

$$|I_1| = |\Lambda_1| = \max(v_+(A), v_+(B))$$

and

$$|I_2| + |\Lambda_2| = v_0(A \cup B).$$

It follows that:

$$\begin{array}{rcll} \text{rank}(S(A,B)) & = & \text{rank}(S(A,B)/\mathcal{I}) & \text{(Lemma 2.74)} \\ & = & \text{rank}(T) + |I_2| + |\Lambda_2| & \text{(Theorem 2.68)} \\ & = & (|I_1| + 1) + |I_2| + |\Lambda_2| & \\ & = & \max(v_+(A), v_+(B)) + v_0(A \cup B) + 1 & \end{array}$$

as required.

Case 3: $\mathcal{MD}(A \cup B) \geq 2$. In this case P has the form:

$$\begin{array}{cc} & \begin{array}{cc} I_1 & I_2 \end{array} \\ \begin{array}{c} \Lambda_1 \\ \\ \Lambda_2 \end{array} & \left(\begin{array}{c|c} Q & \\ \hline & \end{array}\right) \end{array}$$

where Q is regular, non-diagonal and every entry outside of Q is equal to zero. As before, let $T = \mathcal{M}^0[G; I_1, \Lambda_1; Q]$. Since Q is not diagonal the number of components that Q has, when decomposed as in Proposition 2.58, must be strictly less than $\max(|I_1|, |\Lambda_1|)$. Also since $r > 2$ it follows that $\text{rank}(G) = \text{rank}(S_r) = 2$ and so $r_{\min} \leq \text{rank}(S_r) = 2$. We conclude that:

$$r_{\min} + k - 1 \leq 2 + k - 1 \leq \max(|I_1|, |\Lambda_1|).$$

It follows that $\text{rank}(T) = \max(|I_1|, |\Lambda_1|)$. We conclude:

$$\begin{aligned}
\text{rank}(S(A,B)) &= \text{rank}(S(A,B)/\mathcal{I}) && \text{(Lemma 2.74)}\\
&= \text{rank}(T) + |I_2| + |\Lambda_2| && \text{(Theorem 2.68)}\\
&= \max(|I_1|, |\Lambda_1|) + |I_2| + |\Lambda_2| \\
&= \max(v_+(A), v_+(B)) + v_0(A \cup B)
\end{aligned}$$

as required. □

Note that the result is slightly different for $r = 2$ since S_2 is cyclic and so has rank 1, not 2.

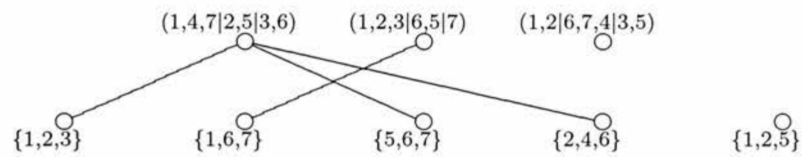
Example 2.76. Let $n = 7$ and $r = 3$ and define the set of images:

$$A = \{\{1,2,3\}, \{1,6,7\}, \{5,6,7\}, \{2,4,6\}, \{1,2,5\}\}$$

and set of partitions:

$$B = \{(1,4,7|2,5|3,6), (1,2,3|6,5|7), (1,2|6,7,4|3,5)\}.$$

Then the graph $\Gamma(A, B)$ is isomorphic to:



which has two isolated vertices so that $v_0(A \cup B) = 2$, $v_+(A) = 2$, $v_+(B) = 4$ and maximum degree $\mathcal{MD}(A \cup B) = 3$. Therefore:

$$\text{rank}(S(A,B)) = \max(2,4) + 2 = 6.$$

Corollary 2.77. *If $S(A,B)/\mathcal{I}$ is regular and has at least two idempotents in one $\mathcal{R}$- or $\mathcal{L}$-class then*

$$\text{rank}(S(A,B)) = \max(|A|, |B|).$$

In particular the semigroups $K(n,r)$ satisfy the conditions of the above Corollary.

2.13.1 Ranks of regular subsemigroups of the full transformation semigroup

In [16], while considering the question of the *length* of finite groups, the authors give a sketch of the proof of the following result, due originally to P. Neumann.

Theorem 2.78 (P. Neumann). $\mathrm{rank}(G) \leq \max(2, \lfloor n/2 \rfloor)$ *for all* $G \leq S_n$.

A similar result for subsemigroups of the full transformation semigroup would be of interest. In general this still seems like a difficult problem:

Open Problem 1. Determine a formula for $\max\{\mathrm{rank}(S) : S \leq T_n\}$.

One feels that the answer, if there is any reasonable answer, will be given by finding a very large block of non-group $\mathcal{H}$-classes in some $\mathcal{J}$-class and taking this set of elements as generators. If we add a number of hypotheses we are able to obtain a positive result of the the above type.

Theorem 2.79. *Let* $n \geq 4$ *and let* $1 < r < n$. *Every regular subsemigroup of* T_n *that is generated by mappings all with rank* r, *and has a unique maximal* $\mathcal{J}$*-class, is generated by at most* $S(n, r)$ *elements. Moreover, the bound is attained by the semigroup* $K(n, r)$.

Proof. Let S be a regular subsemigroup of T_n generated by mappings of rank r and with a unique maximal $\mathcal{J}$-class. Let J_M be the unique maximal $\mathcal{J}$-class of S. Then $\mathrm{rank}(S) = \mathrm{rank}({J_M}^*)$ where the principal factor ${J_M}^*$ is isomorphic to a completely 0-simple $\mathcal{M}^0[G; I, \Lambda; P]$ where $G \leq S_r$ and the matrix P has at most $\binom{n}{r}$ connected components. By Theorem 2.78, since $G \leq S_r$, it follows that $\mathrm{rank}(G) \leq \lfloor r/2 \rfloor$. By Theorem 2.8

$$\mathrm{rank}(S) \leq \max(S(n, r), \binom{n}{r}, \lfloor r/2 \rfloor + \binom{n}{r} - 1).$$

Of course,

$$\max(S(n, r), \binom{n}{r}, \lfloor r/2 \rfloor + \binom{n}{r} - 1) = \max(S(n, r), \lfloor r/2 \rfloor + \binom{n}{r} - 1).$$

Also, for $n \geq 4$ and $1 < r < n$ we have:

$$S(n, r) > \lfloor r/2 \rfloor + \binom{n}{r} - 1.$$

It follows that $\mathrm{rank}(S) \leq S(n, r)$. □

Can we remove the hypothesis that the subsemigroup must have a unique maximal $\mathcal{J}$-class?

Open Problem 2. Let $n \geq 4$ and $1 < r < n$. Prove that any regular subsemigroup of T_n that is generated by mappings of rank r is generated by at most $S(n,r)$ elements.

Another direction might be to consider inverse subsemigroups of the symmetric inverse semigroup I_n.

Open Problem 3. Find $\max\{\text{rank}(S) : S \leq I_n\}$ where S is an inverse subsemigroup of I_n.

Chapter 3

Idempotent generating sets and Hall's marriage theorem

3.1 Semigroups generated by idempotents

In [57, Section 6], where various classes of regular semigroup are being discussed, there is a section devoted to, so called, *semibands*. A semiband is a regular semigroup that is generated by its idempotents. The canonical example of such a semigroup is the semigroup Sing_n of all non-invertible elements of T_n. More generally, the semigroup $K(n,r)$ is a semiband for $r = 1, \ldots, n-1$ (see Proposition 2.14). The reason Howie gives for including a section on semibands in his monograph is because

> "...of the frequency with which such semigroups occur 'in nature' and in the universal property they possess."

The universal property being that every semigroup is embeddable in a semiband. As he points out, many naturally occurring semigroups turn out to be idempotent generated. The proper ideals of the full transformation semigroup, we have already seen, are idempotent generated. The ideals of the semigroup O_n of order preserving mappings were shown to be idempotent generated by Aĭzenštat in [4]. This was reproven in [38] where, in addition, the semigroup of partial order preserving transformations PO_n was also shown to be idempotent generated. A subsemigroup S of T_n is called S_n*-normal* if S is stable under conjugation by elements of S_n (i.e. $\forall \alpha \in S_n : \alpha S \alpha^{-1} = S$). In [65] it is shown that S_n-normal semigroups are idempotent generated which generalizes the result for $K(n,r)$ since $K(n,r)$ is an S_n-normal subsemigroup of T_n.

In addition to these transformation monoids, various linear semigroups have been considered. The proper ideals of the semigroup of all linear transformations of a finite vector space were shown to be semibands in [25] and [29]. Also, the proper ideals of the semigroup of affine mappings of a finite vector space were shown to be idempotent generated in [24]. In [31] the semigroup of $n \times n$ matrices over $\mathbb{Z}$ was shown to be idempotent generated.

It follows from [57, Exercise 6.12] that a finite semigroup is a semiband if and only if every principal factor of the semigroup is idempotent generated. Each of these factors is completely 0-simple and, using the normal form introduced in Section 2.10, we can determine exactly when a completely 0-simple semigroup will be idempotent generated.

Theorem 3.1. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a finite completely* 0*-simple semigroup in Graham normal form. Then S is idempotent generated if and only if S is connected, and the group G is generated by the non-zero entries in the matrix P.*

Proof. Follows from Theorem 2.58. □

Given an idempotent generated semigroup S its *idempotent rank* is defined to be the minimal cardinality of a generating set of idempotents for S. In other words

$$\text{idrank}(S) = \min\{|B| : B \subseteq E(S), \langle B \rangle = S\}.$$

Clearly the idempotent rank is at least as large as the rank. In general, however, they will not be equal. On the other hand, a fair number of semigroups are known to have idempotent rank equal to rank. In [59], the semigroup $K(n,r)$ was shown to have this property. Since then, many other idempotent generated semigroups have been shown to have this property. A list of some other examples, along with their ranks and idempotent ranks, is given in Table 3.1. As with the question of rank, for many of these examples the question of idempotent rank reduces to the same question for a corresponding completely 0-simple semigroup.

Amongst the collection of papers where idempotent rank is considered is a series of papers [68], [67] and [66], where the authors consider semigroups generated by transformations of prescribed partition type. These semigroups are closely related to the S_n-normal semigroups defined above. Stealing the expression from [67], if a semigroup S is idempotent generated and satisfies $\text{rank}(S) = \text{idrank}(S)$ then we say it has an *extremal idempotent generating set.* In [67] it is shown that S_n-normal semigroups have extremal idempotent generating sets. The methods used involve relating the idempotent rank problem to the, so called, Partition Type Conjecture. The Partition Type Conjecture is a generalization of the famous Middle Level Conjecture, attributed to Paul Erdös (see [85]), which asks whether the bipartite graph given by the middle two levels of the poset of all subsets of X_{2n+1}, ordered by inclusion, is Hamiltonian.

It is the connections between finding extremal idempotent generating sets and combinatorial problems in bipartite graphs which will be the main focus of this chapter. Finite idempotent generated completely 0-simple semigroups will be considered with a view to classifying those that have extremal idempotent generating sets. The results will then be applied to a number of concrete examples. In particular, it will be shown that while the existence of Hamiltonian cycle is sufficient for finding extremal idempotent generating sets, it is actually not necessary.

In §3.2 relationships between ranks and idempotent ranks of rectangular 0-bands and their substructures are explored. Then in §3.3 and §3.4 the ranks and idempotent ranks of completely 0-simple semigroups are considered, and the main results of the chapter are given. Some applications of these results are given in §3.5. Nilpotent rank is the subject of §3.6 and in §3.7 the connections between idempotent rank and the problem of counting the number of bases of a semigroup are explored.

3.2 Division and direct products

We now introduce a few notions concerning completely 0-simple semigroups which we will need in what follows. A rectangular band with a zero adjoined is a band (every element is idempotent) and therefore has an extremal idempotent generating set. Recall that we call two $m \times n$ matrices A and B over $\{0,1\}$ *equivalent* if B can be obtained from A by permuting its rows and columns. Given a regular matrix Q with entries in $\{0,1\}$ we will use $ZB(Q)$ to denote the rectangular 0-band with structure matrix Q.

We say that the rectangular 0-band $ZB(P)$ *divides* the rectangular 0-band $ZB(Q)$ if P is equivalent to a submatrix of Q. We say that $ZB(P)$ is an $\mathcal{R}$*-class (respectively* $\mathcal{L}$*-class) filling subsemigroup* of $ZB(Q)$ if $ZB(P)$ divides $ZB(Q)$ and $ZB(P)$ has the same number of non-zero $\mathcal{R}$-classes (respectively $\mathcal{L}$-classes) as $ZB(Q)$.

We call a completely 0-simple semigroup S *tall* if it has at least as many $\mathcal{R}$-classes as $\mathcal{L}$-classes and *wide* if it has no more $\mathcal{R}$-classes than $\mathcal{L}$-classes. Note that S is both tall and wide if and only if the number of $\mathcal{R}$-classes equals the number of $\mathcal{L}$-classes. In this case we say that S is *square*.

Example 3.2. Let S and T be the rectangular 0-bands with structure matrices P and Q, respectively, where:

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Then S is wide, T is tall and T is an $\mathcal{R}$-class filling subsemigroup of S.

The next lemma tells us that finding a suitable submatrix of P is enough to conclude that $ZB(P)$ has an extremal idempotent generating set.

Lemma 3.3. *Let S and T be rectangular 0-bands such that T has an extremal idempotent generating set. Then:*

(i) if S and T are both tall and T is an $\mathcal{L}$-class filling subsemigroup of S then S has an extremal idempotent generating set;

(ii) if S and T are both wide and T is an $\mathcal{R}$-class filling subsemigroup of S then S has an extremal idempotent generating set.

Proof. (i) Let $B \subseteq T \leq S$ be an extremal idempotent generating set for T. If $S = T$ then B is an extremal idempotent generating set for S and we are done. Suppose otherwise so that S is strictly taller than T. Since S and T are both tall, and T is an $\mathcal{L}$-class filling subsemigroup of S, the set B intersects every $\mathcal{L}$-class of S but does not intersect every $\mathcal{R}$-class. Let C be a subset of S, consisting entirely of idempotents, that forms a transversal of the $\mathcal{R}$-classes of S that the set B does not intersect. Such a set C exists since S is regular and so every $\mathcal{R}$-class contains at least one idempotent. We claim that the set $B \cup C$ is an extremal idempotent generating set of S. The number of elements in $B \cup C$ is equal to the number of $\mathcal{R}$-classes in S and since S is tall this number equals $\operatorname{rank}(S)$ by Lemma 2.10. It follows that $|B \cup C| = \operatorname{rank}(S)$ which is the required size for an extremal idempotent generating set. Also, by construction it follows that every element in $B \cup C$ is an idempotent. We are just left to prove that $B \cup C$ generates S. Let $s \in S$. If $s \in T \leq S$ then $s \in \langle B \rangle \subseteq \langle B \cup C \rangle$. Otherwise we have $s \in S \setminus T$ and in this case we let $\{u\} = R_s \cap C$ and let e be some idempotent in T that is $\mathcal{L}$-related to u. Such an element e exists since the subsemigroup T is regular. Finally let $\{v\} = R_e \cap L_s \subseteq T$. Since B generates T and since $v \in T$ it follows that $v \in \langle B \rangle$. Also by definition we have $u \in C$. But now $s = uv$ (since e is an idempotent) and so $s \in \langle B \cup C \rangle$. Since s was arbitrary it follows that $\langle B \cup C \rangle = S$. (ii) Is proved using a dual argument. □

In general the direct product of two completely 0-simple semigroups is not a completely 0-simple semigroup.

Definition 3.4. Let S and T be semigroups each with a zero element. Let

$$Z = \{(x, 0) : x \in S\} \cup \{(0, y) : y \in T\} \subseteq S \times T$$

noting that Z is a two-sided ideal of $S \times T$. Let $S \times_0 T$ denote the Rees quotient $(S \times T)/Z$. We call $S \times_0 T$ the direct product of S and T *with amalgamated zero.*

If S and T are completely 0-simple semigroups then so is $S \times_0 T$. In the particular case where S and T are rectangular 0-bands then so is $S \times_0 T$ and the structure matrix of $S \times_0 T$ is constructed from the structure matrices of S and T in the following way.

Given two matrices P and Q over $\{0,1\}$ the *tensor product* of P and Q, written $P \otimes Q$, is given by:

$$P \otimes Q = \begin{pmatrix} p_{11}Q & \dots & p_{1n}Q \\ \vdots & \ddots & \vdots \\ p_{m1}Q & \dots & p_{mn}Q \end{pmatrix}$$

with 0 and 1 multiplying in the usual way. For any two regular matrices P and Q over $\{0,1\}$ the matrix $P \otimes Q$ is regular and $ZB(P) \times_0 ZB(Q) \cong ZB(P \otimes Q)$.

The following simple lemma provides an upper bound for the idempotent rank of the direct product, with amalgamated zero, of two semigroups.

Lemma 3.5. *Let S and T be idempotent generated semigroups each with a zero element. Then*

$$\mathrm{idrank}(S \times_0 T) \leq \mathrm{idrank}(S)\mathrm{idrank}(T).$$

Proof. First note that if S and T are idempotent generated then so is $S \times T$ and $\mathrm{idrank}(S \times T) \leq \mathrm{idrank}(S)\mathrm{idrank}(T)$. Indeed, if A and B are idempotent generating sets for S and T, respectively, then we claim that $A \times B \subseteq S \times T$ is an idempotent generating set for $S \times T$. To see this let $(s,t) \in S \times T$ and write $s = e_1 \cdots e_k$ where $e_i \in A$ for $1 \leq i \leq k$ and $t = f_1 \cdots f_r$ where $f_i \in B$ for all $1 \leq i \leq r$. Suppose without loss of generality that $k \leq r$. Then we have:

$$(s,t) = (e_1, f_1) \cdots (e_{k-1}, f_{k-1})(e_k, f_k)(e_k, f_{k+1})(e_k, f_{k+2}) \cdots (e_k, f_r) \in \langle A \times B \rangle$$

so the set $A \times B$ generates $S \times T$. Also we have $|A \times B| = |A||B|$ so the set $A \times B$ has the required size. Using the claim we have:

$$\mathrm{idrank}(S \times_0 T) = \mathrm{idrank}((S \times T)/Z) \leq \mathrm{idrank}(S \times T) \leq \mathrm{idrank}(S)\mathrm{idrank}(T)$$

as required. □

From the above lemma we now give a sufficient condition for $S \times_0 T$ to have an extremal idempotent generating set.

Lemma 3.6. *Let S and T be rectangular 0-bands such that* $\mathrm{idrank}(S) = \mathrm{rank}(S)$ *and* $\mathrm{idrank}(T) = \mathrm{rank}(T)$. *If S and T are both tall (or both wide) then*

$$\mathrm{idrank}(S \times_0 T) = \mathrm{rank}(S \times_0 T).$$

Proof. Let P and Q be the structure matrices of S and T, respectively, where P is an $a \times b$ matrix and Q is a $c \times d$ matrix. Since S and T are both tall it follows that $b \geq a$ and $d \geq c$. We have

$$\operatorname{idrank}(S \times_0 T) \geq \operatorname{rank}(S \times_0 T) = \max\{bd, ac\} = bd.$$

Also, by Lemma 3.5, we have

$$\operatorname{idrank}(S \times_0 T) \quad \leq \quad \operatorname{idrank}(S)\operatorname{idrank}(T) = \operatorname{rank}(S)\operatorname{rank}(T) = bd.$$

We conclude that

$$\operatorname{idrank}(S \times_0 T) = bd = \operatorname{rank}(S \times_0 T).$$

□

The condition that S and T are either both tall or both wide is necessary as the following example demonstrates.

Example 3.7. Let S and T be rectangular 0-bands with structure matrices P and Q, respectively, where

$$P = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Then $S \otimes T$ has structure matrix

$$P \otimes Q = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

In this example S and T both have extremal idempotent generating sets but $S \otimes T$ does not have one. Indeed, T is a band and so obviously has an extremal idempotent generating set, and S is generated by $\{(1,1),(2,2),(3,2),(4,2)\}$. The semigroup $S \otimes T$ does not have an extremal idempotent generating set because it does not satisfy SHC (see the next section for the definition of SHC and the proof of this fact).

Example 3.8. Let S and T be the 4-element right zero semigroup and 3-element left zero semigroup, respectively (both with zero adjoined). Then $S \otimes T$ is the 3×4 rectangular band (with 0 adjoined) which does have an extremal idempotent generating set even though S is tall and T is wide.

Finally we include an example that shows $S \otimes T$ can have an extremal idempotent generating set even if neither S nor T does.

Example 3.9. Let S and T be rectangular 0-bands with structure matrices P and Q, respectively, where

$$P = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Then $S \otimes T$ has structure matrix

$$P \otimes Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

and in this example neither S nor T both have extremal idempotent generating sets but $S \otimes T$ does have an extremal idempotent generating set. For example the set $A = \{(1,1),(2,2),(3,2),(4,1),(5,3),(6,4)\} \subseteq E(S \otimes T)$ generates $S \otimes T$.

3.3 Idempotent generated completely 0-simple semigroups

Recall that with every completely 0-simple semigroup S we may associate a rectangular 0-band $T = S\natural$ with the same dimensions. In the previous chapter we saw a connection between the generating sets of T and those of S. In the case when S is idempotent generated this connection is even stronger.

Lemma 3.10. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a finite idempotent generated completely 0-simple semigroup and let $T = S\natural$ be the natural rectangular 0-band homomorphic image of S. Then $A \subseteq S$ is a generating set for S if and only if $A\natural \subseteq T$ is a generating set for T.*

Proof. If A generates S then, since $\natural$ is a homomorphism, it follows that $A\natural$ generates T. For the converse, if $A\natural$ generates T then it follows that A generates at least one element of every non-zero $\mathcal{H}$-class of S. But S is idempotent generated so $\langle A \rangle = S$ by Lemma 2.4. □

Combined with the results for rectangular 0-bands in the previous chapter we have:

Lemma 3.11. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be an idempotent generated completely 0-simple semigroup and let $T = S\natural$ be the natural rectangular 0-band homomorphic image of S. Then:*

(i) $\text{rank}(S) = \text{rank}(T) = \max(|I|, |\Lambda|)$;

(ii) $\text{idrank}(S) = \text{idrank}(T)$.

In particular S has an extremal idempotent generating set if and only if T has one.

Proof. (i) See Theorem 2.13. (ii) If $X \subseteq E(S)$ generates S then $X\natural \subseteq E(T)$ with $|X\natural| \leq |X|$ and $\langle X\natural \rangle = T$. Therefore $\text{idrank}(S) \geq \text{idrank}(T)$. For the converse let $Y \subseteq E(T)$ generate T. Then $X = \{(i, p_{\lambda i}^{-1}, \lambda) : (i, \lambda) \in Y\}$ is a subset of the non-zero idempotents of S satisfying $X\natural = Y$ and so generates S by Lemma 3.10. Since $|X| = |Y|$ it follows that $\text{idrank}(S) \leq \text{idrank}(T)$.

□

Given a rectangular 0-band S we will often suppose that the index sets consist of natural numbers (i.e. that $I = \{1, \ldots, m\}$ and $\Lambda = \{1, \ldots, n\}$). This is a possible source of confusion since we will still want to view I and Λ as being disjoint sets. In some situations, in order to avoid ambiguity, we will use $I = \{i_1, \ldots, i_m\}$ and $\Lambda = \{\lambda_1, \ldots, \lambda_n\}$.

Recall the correspondence between non-zero products of idempotents in $T = S\natural$ and paths between vertices of I and those of Λ in the graph $\Delta(P)$. Indeed, the equality $(i_1, \lambda_1) \cdots (i_k, \lambda_k) = (i_1, \lambda_k)$, with $(i_l, \lambda_l) \in E(T)$ for $1 \leq l \leq k$, holds in T precisely when $p_{\lambda_1 i_2}, \ldots, p_{\lambda_{k-1} i_k}$ are all non-zero, which is the same as saying that $i_1 \to \lambda_1 \to \ldots \to i_k \to \lambda_k$ is a path in the graph $\Delta(P)$.

Lemma 3.12. *The set $E' \subseteq E(T)$ of idempotents generates $(i_1, \lambda_k) \in T$ if and only if there is a path $i_1 \to \lambda_1 \to i_2 \to \lambda_2 \to \ldots \to i_k \to \lambda_k$ in $\Delta(P)$ such that $(i_m, \lambda_m) \in E'$ for $1 \leq m \leq k$.* □

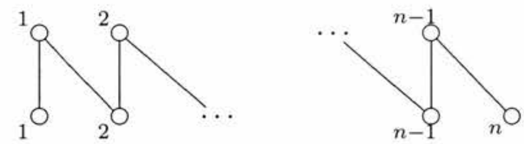
There is a limited range of values that the idempotent rank can take.

Proposition 3.13. *If $S = \mathcal{M}^0[G; I, \Lambda; P]$ is idempotent generated then*

$$\max(|I|, |\Lambda|) \leq \text{idrank}(S) \leq |I| + |\Lambda| - 1.$$

Proof. The first inequality holds since $\text{idrank}(S) \geq \text{rank}(S) = \max(|I|, |\Lambda|)$ by Lemma 2.3. As for the second inequality, since S is idempotent generated the graph $\Delta(P)$ is connected. Let $\mathcal{T}$ be a spanning tree of $\Delta(P)$ noting that $|\mathcal{T}| = |I| + |\Lambda| - 1$. The set of idempotents given by the edges in this tree is a generating set for S as a consequence of Lemma 3.12. □

Example 3.14. It was observed in [4] that the semigroup O_n of all non-identity order preserving mappings of X_n is idempotent generated. Let S be the principal factor of the unique maximal $\mathcal{J}$-class of O_n. The semigroup S has $n-1$ non-zero $\mathcal{R}$-classes and n non-zero $\mathcal{L}$-classes. The graph $\Delta(S)$ is isomorphic to

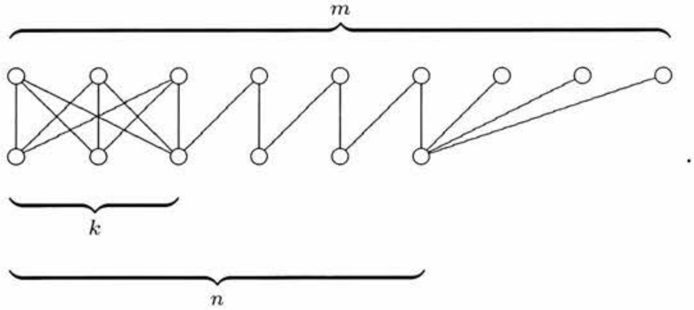


which is connected but has the property that the removal of any edge will disconnect the graph. As a result, every idempotent is required in order to generate the completely 0-simple semigroup S and

$$\text{idrank}(S) = n + (n-1) - 1 = 2n - 2.$$

In fact, with $|I|$ and $|\Lambda|$ fixed, for every j in the range $\max(|I|, |\Lambda|)$ up to $|I| + |\Lambda| - 1$ there are semigroups with idempotent rank equal to j.

Example 3.15. Let $|I| = m$ and $|\Lambda| = n$. Suppose without loss of generality that $m \geq n$. For $1 \leq k \leq n$ draw a complete bipartite graph $K_{k,k}$ and then extend it using a tree to obtain the following graph



Let S be the rectangular 0-band with structure matrix P where $\Delta(P)$ is isomorphic to the above graph. Then

$$\text{idrank}(S) = k + 2(n-k) + (m-n) = n + m - k$$

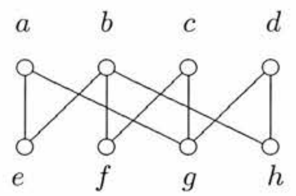
which ranges between $\max(m, n)$ and $n + m - 1$ as $1 \leq k \leq n$.

We now return our attention to the question of which completely 0-simple semigroups have extremal idempotent generating sets. We start by considering square idempotent generated Rees matrix semigroups (i.e. when $|I| = |\Lambda|$) and then extend our results to deal with non-square ones. Note that if S is a square idempotent generated Rees matrix semigroup, with structure matrix P, then $\Delta(P)$ is a connected and balanced bipartite graph.

If $\Delta(P)$ is Hamiltonian then S will have an extremal idempotent generating set. Indeed if $i_1 \to \lambda_1 \to \ldots \to i_n \to \lambda_n \to i_1$, with $n = |I| = |\Lambda|$, is a Hamiltonian cycle then, by Lemma 3.12, the subset $E' = \{(i_j, \lambda_j) : 1 \leq j \leq n\}$ of $E(S)$ generates $T = S\natural$. It is, however, not necessary that $\Delta(P)$ is Hamiltonian in order for S to have an extremal idempotent generating set.

Example 3.16. Let S be the rectangular 0-band with structure matrix

$$P = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$



.

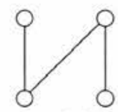
$\Delta(P)$

The graph $\Delta(P)$ is not Hamiltonian since the edges $\{a, e\}$, $\{e, b\}$, $\{b, h\}$, $\{h, d\}$, $\{d, g\}$ and $\{g, a\}$ must all be included in any Hamiltonian cycle and they themselves already form a closed path. However, the set of idempotents $E' = \{(1, 1), (2, 2), (3, 3), (4, 4)\}$ generates S and so S has an extremal idempotent generating set.

On the other hand, it is clear that for a square idempotent generated Rees matrix semigroup S to have an extremal idempotent generating set the graph $\Delta(P)$ must have a perfect matching. Indeed, given a generating set consisting of n idempotents the edges in $\Delta(P)$ corresponding to these elements will constitute a perfect matching in $\Delta(P)$. This is equivalent, as a consequence of Hall's marriage theorem, to saying that in $\Delta(P)$ for every subset X of I we have $|N(X)| \geq |X|$. It is, however, not sufficient that $\Delta(P)$ has a perfect matching (and is connected) to conclude that S has an extremal idempotent generating set.

Example 3.17. Let S be the rectangular 0-band with structure matrix

$$P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$



$\Delta(P)$

The graph $\Delta(P)$ is connected and has a perfect matching but $\text{idrank}(S) = 3$ (for the same reason as Example 3.14) while $\text{rank}(S) = 2$.

Definition 3.18. We call a subset A of the semigroup $S = \mathcal{M}^0[G; I, \Lambda; P]$ a *sparse cover* of S if $|A| = \max(|I|, |\Lambda|)$ and A intersects every non-zero $\mathcal{R}$-class and every non-zero $\mathcal{L}$-class of S.

Theorem 2.10 in Section 2.3 says that every rectangular 0-band, and as a consequence every idempotent generated completely 0-simple semigroup, has at least one sparse cover that generates it. Of course, when $|I| = |\Lambda| = n$ the semigroup $S = \mathcal{M}^0[G; I, \Lambda; P]$ has at most $|G|^n n!$ generating sets of size n, and this number is attained precisely when any sparse cover of S generates S.

Lemma 3.19. *Let $\Gamma = I \cup \Lambda$ be a connected and balanced bipartite graph. Then the following are equivalent:*

(i) $|N(X)| > |X|$ *for every non-empty proper subset X of I;*

(ii) $|N(Y)| > |Y|$ *for every non-empty proper subset Y of Λ.*

Proof. Let X be a non-empty proper subset of I. Suppose that $|N(X)| \leq |X|$. It follows that $(|\Lambda| - |N(X)|) - |\Lambda| \geq (|I| - |X|) - |I|$ and so $|\Lambda \setminus N(X)| \geq |I \setminus X| \geq |N(\Lambda \setminus N(X))|$. □

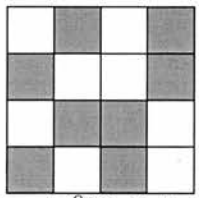
If a connected balanced bipartite graph $\Gamma = I \cup \Lambda$ satisfies either, and hence both, of the conditions given in Lemma 3.19 we say that Γ satisfies *strong Hall's condition* (SHC for short). We say that $S = \mathcal{M}^0[G; I, \Lambda; P]$, with $|I| = |\Lambda|$, satisfies SHC if the graph $\Delta(P)$ does (see Figure 3.1 for an example of this).

We now describe the class of square idempotent generated completely 0-simple semigroups that have extremal idempotent generating sets.

Theorem 3.20. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be an idempotent generated completely 0-simple semigroup with $|I| = |\Lambda|$. Then the following are equivalent:*

(i) $\text{rank}(S) = \text{idrank}(S)$*;*

(ii) any sparse cover of S generates S;



$S = \mathcal{M}^0[G; I, \Lambda; P]$

$$Q = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

$T = \mathcal{M}^0[\{1\}; I, \Lambda; Q]$

$\Delta(P)$

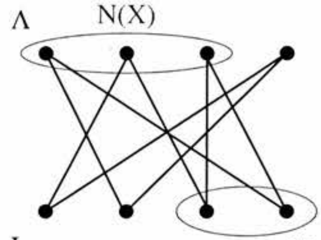


Figure 3.1: Egg-box picture of a Rees matrix semigroup S, the structure matrix Q of $T = S\natural$, the associated graph $\Delta(P)$ and the neighbourhood of a collection of vertices X. The set X satisfies $|N(X)| > |X|$. In fact, in this graph for every proper subset Y of I we have $|N(Y)| > |Y|$ and so S satisfies SHC.

(iii) S satisfies SHC.

Proof. As a consequence of Lemma 3.11 it is sufficient to prove the result for rectangular 0-bands. Let $S = (I \times \Lambda) \cup \{0\}$ be an $n \times n$ rectangular 0-band with $I = \Lambda = \{1, \dots, n\}$. Condition (ii) for rectangular 0-bands says that for every $\beta \in S_n$ (the symmetric group on $\{1, \dots, n\}$) the set $X(\beta) = \{(1, 1\beta), \dots, (n, n\beta)\}$ generates S.

((i) $\Rightarrow$ (ii)) Suppose without loss of generality that $E' = \{(1,1), \dots, (n,n)\}$ is an extremal idempotent generating set for S. Let $X(\beta) = \{(1, 1\beta), \dots, (n, n\beta)\}$ for some $\beta \in S_n$. Let k be the order of β in S_n. Then

$$(i, i) = (i, i\beta)(i\beta, i\beta^2)\dots(i\beta^{k-1}, i\beta^k)$$

where each of the terms on the right hand side belongs to $X(\beta)$. Therefore $\langle X(\beta)\rangle \supseteq \langle E'\rangle = S$ and we conclude that $X(\beta)$ generates S.

((ii) $\Rightarrow$ (iii)) Suppose that S does not satisfy SHC and let J be a non-empty proper subset of I such that $|N(J)| \le |J|$. We will find a sparse cover of S that does not generate S. Let $I = \{i_1, \dots, i_n\}$, $\Lambda = \{\lambda_1, \dots, \lambda_n\}$ with $J = \{i_1, \dots, i_q\}$ and $N(J) = \{\lambda_1, \dots, \lambda_l\}$ where $l \le q$. We claim that the sparse cover $Y =$

$\{(i_1, \lambda_1), \ldots, (i_n, \lambda_n)\}$ does not generate S, in particular $(i_n, \lambda_1) \notin \langle Y \rangle$. Consider a path $i_n \to \lambda_n \to i_{q_1} \to \lambda_{q_1} \to \ldots \to i_{q_t} \to \lambda_{q_t}$, corresponding to a product of elements of Y, starting at i_n. We know that $N(\Lambda \setminus N(J)) \subseteq I \setminus J$ which implies that $i_{q_1} \in I \setminus J$. It follows, since $l \leq q$, that $\lambda_{q_1} \in \Lambda \setminus N(J)$ and so $i_{q_2} \in I \setminus J$. Continuing in this way we see that $i_{q_t} \in I \setminus J$ giving $\lambda_{q_t} \in \Lambda \setminus N(J)$ and as a consequence $\lambda_{q_t} \neq \lambda_1$.

((iii) $\Rightarrow$ (i)) Suppose that S satisfies SHC. Then in particular S satisfies HC and, by Hall's marriage theorem, the graph $\Delta(P)$ has a perfect matching $\pi : I \to \Lambda$. We claim that $M = \{(i, \pi(i)) : i \in I\}$ generates S. Let $(i, \lambda) \in S$ be arbitrary. We will construct a path from i to λ in $\Delta(P)$ such that the first, third, fifth, and so on, edges all belong to belong to M. Start with the path $i \to \pi(i)$. If $\pi(i) = \lambda$ we are done. Otherwise by SHC we know that $|N(\pi(i))| > 1$ which allows us to choose $i_2 \in I \setminus \{i\}$ and extend our path to $i \to \pi(i) \to i_2 \to \pi(i_2)$. If $\pi(i_2) = \lambda$ we are done. Otherwise by SHC we know that $|N(\{\pi(i), \pi(i_2)\})| > 2$ and so there exists an $i_3 \in I \setminus \{i, i_2\}$ which we can extend the path to. Continuing in this way since all of the i_r are distinct and π is a perfect matching, all the $\pi(i_r)$ are distinct. Since $|I| = |\Lambda|$ is finite, we will eventually have a path from i to λ corresponding to a product of elements of M. □

The above result can be extended to give the general result for the non-square case. This will be done in Theorem 3.27. First we note that if we can find a 'subsquare' that satisfies SHC then this is sufficient for proving the existence of an extremal idempotent generating set.

Proposition 3.21. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be an idempotent generated completely 0-simple semigroup. If $\Delta(P)$ has a connected and balanced, bipartite subgraph on $2\min(|I|, |\Lambda|)$ vertices that satisfies SHC then*

$$\text{idrank}(S) = \text{rank}(S) = \max(|I|, |\Lambda|).$$

Proof. Without loss of generality assume that $|I| \leq |\Lambda|$. As a consequence of Lemma 3.11 it is sufficient to prove the result for rectangular 0-bands. The result now follows from Lemma 3.3 and Theorem 3.20. □

The converse of the above result does not hold. It is not true that if a completely 0-simple semigroup has an extremal idempotent generating set then we can necessarily find a subsquare that satisfies SHC.

Example 3.22. Let S be the rectangular 0-band with structure matrix

$$P = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

In this example the set $A = \{(1,1),(2,2),(3,3),(4,3)\}$ consists entirely of idempotents and generates S while none of the sub 3×3 rectangular 0-bands of S have idempotent bases (since none of them satisfy SHC).

We now describe a process of extending non-square matrices to square matrices by making copies of existing rows (columns). By a *square extension* of an $m \times n$ matrix P, where $m \leq n$, we mean an $n \times n$ matrix Q where the first m rows are the same as the first m rows of P and each of the remaining rows is the same as one of these first m rows. More precisely we have the following definition.

Definition 3.23. Let $I = \{1, \ldots, n\}$ and $\Lambda = \{1, \ldots, m\}$ with $m \leq n$, and let $P = (p_{\lambda i})$ be an $m \times n$ matrix with entries in $\{0, 1\}$ indexed by Λ and I. For this definition we no longer want to thing of I and Λ as being disjoint sets. In particular we have $I \cap \Lambda = \Lambda = \{1, \ldots, m\}$. Let $F = \{B_\lambda : \lambda \in \Lambda\}$ be a set of disjoint subsets of I such that $\lambda \in B_\lambda$ for all $\lambda \in \Lambda$ and $\bigcup_{\lambda \in \Lambda} B_\lambda = I$. Given such a partition F of I define $\bar{f} : I \to \Lambda$ so that $i \in B_{\bar{f}(i)}$. Note here that $\bar{f}$ is a retraction.

By the square extension of P by F we mean the $n \times n$ matrix Q with entries $q_{xy} = p_{\bar{f}(x)y}$ for $1 \leq x, y \leq n$. We use $\mathrm{Sq}(P)$ to denote the set of all square extensions of the matrix P.

Example 3.24. If

$$P = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

then $\mathrm{Sq}(P) = \{Q_1, Q_2, \ldots, Q_9\}$ where

$$Q_1 = \left(\begin{array}{ccccc} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{array}\right), Q_2 = \left(\begin{array}{ccccc} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ \hline 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{array}\right), Q_3 = \left(\begin{array}{ccccc} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ \hline 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{array}\right),$$

$$Q_4 = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}, Q_5 = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}, Q_6 = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ \hline 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

and $\{Q_7, Q_8, Q_9\}$ is the set of matrices $\{Q_4, Q_5, Q_6\}$ with the last two rows swapped around. In particular if P is a square matrix then $\mathrm{Sq}(P) = \{P\}$.

Given a rectangular 0-band S we also define $\mathrm{Sq}(S)$ to be the set of rectangular 0-bands corresponding to the matrices in the set $\mathrm{Sq}(P)$.

Definition 3.25. A subsemigroup T of a semigroup S is called a *retract* of S if there is a homomorphism $f : S \to T$, called a *retraction*, such that $f(t) = t$ for all $t \in T$.

Lemma 3.26. *Let S be a completely 0-simple semigroup and let $\mathrm{Sq}(S)$ be the set of square extensions of S. For every T in $\mathrm{Sq}(S)$ the semigroup S is a retract of the semigroup T.*

Proof. Follows from the definition of square extension and [57, Section 3.5] on congruences of completely 0-simple semigroups. □

With $I = \{1, \ldots, m\}, \Lambda = \{1, \ldots, n\}, A \subseteq I$ and $b \in \Lambda$ we define

$$(A, b) = \{(a, b) : a \in A\} \subseteq I \times \Lambda.$$

Theorem 3.27. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be an idempotent generated completely 0-simple semigroup. Then S has an extremal idempotent generating set if and only if $\Delta(Q)$ satisfies SHC for some $Q \in \mathrm{Sq}(P)$.*

Proof. As a consequence of Lemma 3.11 it is sufficient to prove the result for rectangular 0-bands. Let $I = \{1, \ldots, n\}$ and $\Lambda = \{1, \ldots, m\}$ with $m \leq n$. Let S be a rectangular 0-band indexed by I and Λ with structure matrix $P = (p_{\lambda i})$ and extremal idempotent generating set

$$B = (A_1, 1) \cup (A_2, 2) \cup \ldots \cup (A_m, m)$$

where $A_i \neq \varnothing$ for $1 \leq i \leq m$ and $\{1, \ldots, n\}$ is the disjoint union of the sets A_i for $1 \leq i \leq m$. Also suppose, without loss of generality, that $j \in A_j$ for $1 \leq j \leq m$. Let Q be the square extension of P by F where $F = \{A_1, \ldots, A_m\}$ (see Figure 3.2). Let T be the rectangular 0-band defined by Q. We claim that T has an extremal idempotent generating set and thus, by Theorem 3.20, the

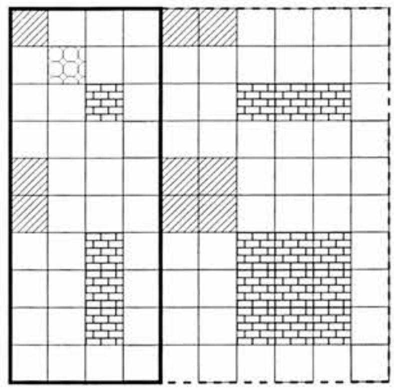


Figure 3.2: Extending a 10×4 semigroup $S = \mathcal{M}^0[G; I, \Lambda; P]$ using $F = \{A_1, A_2, A_3, A_4\}$ where $A_1 \times A_1 =$ ▨, $A_2 \times A_2 =$ ▦, $A_3 \times A_3 =$ ▤, and $A_4 \times A_4 =$ □.

graph $\Delta(Q)$ satisfies SHC. For $1 \leq i \leq m$ choose and fix a bijection $\phi_i : A_i \to A_i$. Since I is equal to the disjoint union $\bigcup_{i=1}^{m} A_i$ we can combine these functions together to give $\phi : I \to I$ where $\phi(x) = \phi_i(x)$ for $x \in A_i$. We claim that the set $X = \{(1, \phi(1)), \ldots, (n, \phi(n))\}$ is a subset of $E(T)$ and that $\langle X \rangle = T$. For the first part note that if $x, y \in A_i$ for some $1 \leq i \leq m$ then

$$q_{xy} = p_{\bar{f}(x)y} = p_{iy} = p_{\bar{f}(y)y} = 1,$$

since $(y, \bar{f}(y))$ belongs to B and is therefore an idempotent. In particular $q_{\phi(j)j} = 1$ for $1 \leq j \leq n$ since j and $\phi(j)$ both belong to the set $A_{\bar{f}(j)}$. It follows from this that $A_j \times A_j$ is a subset of the idempotents of T for $1 \leq j \leq m$. Now we observe that $A_i \times A_i$ is contained in $\langle X \rangle$ for $1 \leq i \leq m$ since given $(x, y) \in A_i \times A_i$ we have

$$(x, y) = (x, \phi_i(x))({\phi_i}^{-1}(y), y) = (x, \phi_i(x))({\phi_i}^{-1}(y), \phi_i({\phi_i}^{-1}(y)))$$

where $q_{{\phi_i}^{-1}(y)\phi_i(x)} = 1$ since $\phi_i(x), {\phi_i}^{-1}(y) \in A_i$. Given any element (x, y) of T the subset $\{y\}$ of $\{1, \ldots, n\}$ can be extended to a transversal J of the sets $A_1, \ldots, A_m$. Consider the subset $I \times J$ of T. By construction $\{I \times J\} \cup \{0\}$ is isomorphic to S and since B generates S it follows that

$$\langle((A_1 \times A_1) \cup \ldots \cup (A_m \times A_m)) \cap (I \times J)\rangle = \{I \times J\} \cup \{0\}.$$

We conclude that

$$(x, y) \in I \times J \subseteq \langle ((A_1 \times A_1) \cup \ldots \cup (A_m \times A_m)) \rangle \subseteq \langle X \rangle,$$

and since (x, y) was arbitrary it follows that X generates T.

For the converse let Q be a square extension of P by $F = \{B_1, \ldots, B_m\}$ where $\Delta(Q)$ satisfies SHC. Let T be the rectangular 0-band defined by Q. By Theorem 3.20 we know that T has an extremal idempotent generating set. Let $\tau \in S_n$ be such that $Y = \{(1, \tau(1)), \ldots, (n, \tau(n))\}$ is an extremal idempotent generating set of T. Define a map $\psi : T \to S$ by $\psi((i, \lambda)) = (i, \bar{f}(\lambda))$ and $\psi(0) = 0$. The map ψ is an onto homomorphism (it is a retraction mapping from T to S) and $|\psi(Y)| = n$. Since $Y \subseteq E(T)$ and ψ is a homomorphism it follows that $\psi(Y) \subseteq E(S)$. Since $\langle Y \rangle = T$ and ψ is a homomorphism it follows that $\langle \psi(Y) \rangle = S$ with $|\psi(Y)| = n$ and therefore S has an extremal idempotent generating set. □

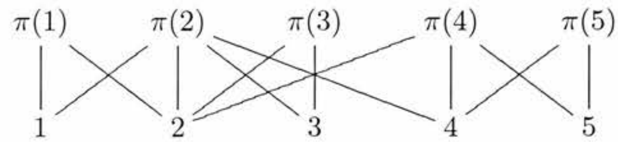
3.4 Regular and symmetric bipartite graphs

The main result of the previous section gives necessary and sufficient conditions for an idempotent generated completely 0-simple semigroup to have an extremal idempotent generating set. Unfortunately, testing SHC directly is a time consuming business since it involves looking at all the subsets of a set. In practice it will be useful for us to note a few conditions which are slightly stronger than SHC but are easier to check.

Lemma 3.28. *If $\Gamma = X \cup Y$ is a k-regular, connected and balanced bipartite graph then Γ satisfies SHC.*

Proof. The number of edges adjacent to the vertices of X and $N(X)$ are $k|X|$ and $k|N(X)|$ respectively. The set of edges adjacent to X is a subset of the edges that are adjacent to $N(X)$ giving $k|N(X)| \geq k|X|$ which implies $|N(X)| \geq |X|$. When $|N(X)| = |X|$ the edges that are adjacent to X are precisely the edges that are adjacent to $N(X)$. It follows that $X \cup N(X)$ is a connected component of Γ which, since Γ is connected, implies that $|X| = |Y|$. □

We say that a balanced bipartite graph $\Gamma = X \cup Y$ with a perfect matching $\pi : X \to Y$ (a bijection such that $\{x, \pi(x)\} \in E(\Gamma)$ for every x in X) has a *symmetric* distribution of edges with respect to the matching π if $d(x) = d(\pi(x))$ for every x in X (See Figure 3.3 for an example). Note that if $\Gamma = X \cup Y$ is



$$
\begin{array}{ll}
d(1) = d(\pi(1)) = 2, & d(2) = d(\pi(2)) = 4, \\
d(3) = d(\pi(3)) = 2, & d(4) = d(\pi(4)) = 3, \\
d(5) = d(\pi(5)) = 2. &
\end{array}
$$

Figure 3.3: A bipartite graph that has a symmetric distribution of edges with respect to the perfect matching π.

k-regular and connected then Γ satisfies HC which implies that Γ has a perfect matching and Γ clearly has a symmetric distribution of idempotents with respect to this matching. Thus Γ being k-regular and connected is a stronger property than Γ having a symmetric distribution of idempotents with respect to some perfect matching.

Lemma 3.29. *If $\Gamma = X \cup Y$ is a connected and balanced bipartite graph that has a symmetric distribution of edges with respect to some perfect matching $\pi : X \to Y$ then Γ satisfies SHC.*

Proof. Let A be a non-empty (not necessarily proper) subset of X. Since $|\pi(A)| = |A|$ and $\pi(A) \subseteq N(A)$ it follows that $|N(A)| \geq |\pi(A)| = |A|$. In addition to this if $|N(A)| = |A|$ then $|N(A)| = |\pi(A)|$ which means that $N(A) = \pi(A)$. But the total number of edges adjacent to A is $\sum_{a \in A} d(a)$ while the total number of edges adjacent to $N(A) = \pi(A)$ is

$$\sum_{b \in \pi(A)} d(b) = \sum_{a \in A} d(\pi(a)) = \sum_{a \in A} d(a).$$

It follows that $A \cup N(A) = A \cup \pi(A)$ is a connected component of $X \cup Y$ but since Γ is connected this means that $A = X$. $\square$

We say that $S = \mathcal{M}^0[G; I, \Lambda; P]$ has a k-uniform distribution of idempotents if the graph $\Delta(P)$ is k-regular. We say that $S = \mathcal{M}^0[G; I, \Lambda; P]$ has a symmetric distribution of idempotents with respect to some perfect matching if $\Delta(P)$ has a symmetric distribution of edges with respect to some perfect matching.

Loosely speaking, the results above tell us that if we can find some symmetry in $\Delta(P)$ (or equivalently in the egg-box diagram of S) then we stand a chance of showing that SHC holds without having to resort to any laborious computation.

3.5 Applications

3.5.1 Ideals of the full transformation semigroup

Our first application of the results of the previous two sections is to prove that the semigroup $K(n,r)$ has an extremal idempotent generating set. The original proof [59] used a fairly complicated Pascal's triangle type induction. Recall that, by Proposition 2.14, the semigroup $K(n,r)$ is generated by elements in its unique maximal $\mathcal{J}$-class.

Lemma 3.30. *Let J_r denote the top $\mathcal{J}$-class of the semigroup $K(n,r)$ and $P_r = K(n,r)/K(n,r-1)$ for $1 < r < n$. Then*

$$\text{rank}(K(n,r)) = \text{rank}(P_r), \qquad \text{idrank}(K(n,r)) = \text{idrank}(P_r).$$

□

Our aim is to prove that the $(\binom{n}{r}, S(n,r))$ bipartite graph $\Delta(P_r)$, with $1 < r < n$, has a balanced $(\binom{n}{r}, \binom{n}{r})$ bipartite subgraph that is connected and has a perfect matching $\pi : X \to Y$ with respect to which P_r has a symmetric distribution of idempotents. Then, as a consequence of Theorem 3.20 and Lemma 3.29, it will follow that $K(n,r)$ has an extremal idempotent generating set.

Let $\mathcal{F}_r$ denote the family of all subsets of $X_n = \{1,\ldots,n\}$ with size r, so that

$$\mathcal{F}_r = \{A \subseteq X_n : |A| = r\}.$$

Define $\mathcal{K}_r \subseteq X_n \times X_n$ to be the family of all partitions of X_n with weight r. By Proposition 2.14 the graph $\Delta(P_r)$ is isomorphic to the graph $\Gamma(\mathcal{F}_r, \mathcal{K}_r)$ of Definition 2.73. It has vertex set $\mathcal{F}_r \cup \mathcal{K}_r$ (disjoint union) with $A \in \mathcal{F}_r$ connected to $K \in \mathcal{K}_r$ if and only if A is a transversal of K. We are mainly concerned with subsets and partitions of $X_n = \{1,\ldots,n\}$ and will want to use modular arithmetic on these symbols rather than the usual $\{0,\ldots,n-1\}$. In view of this fact we will use the convention that $n \bmod n = n$ rather than $n \bmod n = 0$.

Definition 3.31. Given $a, b \in X_n$ we define:

$$[a,b] = \{a, a+1, a+2, \ldots, b\}$$

with all entries reduced mod n. We call $[a,b]$ the *interval* between a and b.

Example 3.32. Given $2, 4 \in X_5$ we have $[4,2] = \{4,5,1,2\}$ while $[2,4] = \{2,3,4\}$. Also $[1,1] = \{1\}$.

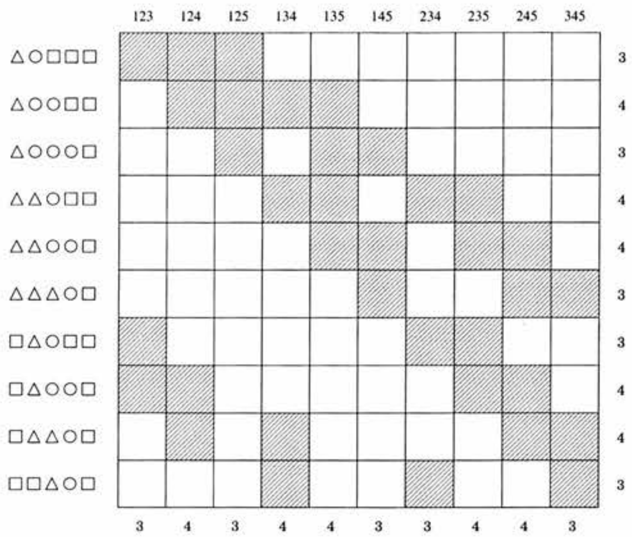


Figure 3.4: The symmetric distribution of idempotents with respect to the perfect matching ϕ corresponding to the idempotents along the main diagonal in J_3 of T_5. The kernel that labels row i is $\phi(I)$ where I is the image that labels column i. The shaded boxes give the positions of the idempotents in J_3.

Definition 3.33. Define $\phi : \mathcal{F}_r \to \mathcal{K}_r$ by:

$$\phi(I) = \{[i_1, i_2 - 1], [i_2, i_3 - 1], \ldots, [i_r, i_1 - 1]\}$$

where $I = \{i_1, \ldots, i_r\} \subseteq X_n$, $i_1 < \ldots < i_r$ and all entries are reduced mod n.

We are using the convention of writing partitions of X_n as sets of sets. For example, if $n = 5$, $r = 3$ and $I = \{2, 3, 4\}$ then

$$\phi(\{2,3,4\}) = \{[2,2],[3,3],[4,1]\} = \{\{2\},\{3\},\{4,5,1\}\}.$$

It is clear from the definition that the map ϕ is injective. In a natural way it associates a kernel, with weight r, to each r-subset of X_n (for example see Figure 3.4). Note that every image I is a transversal of the kernel $\phi(I)$.

Lemma 3.34. *The balanced bipartite graph $\Gamma' = \mathcal{F}_r \cup \phi(\mathcal{F}_r)$ is connected and has a symmetric distribution of edges with respect to the perfect matching ϕ.*

Proof. First we will show that the graph is connected. By definition, each $A \in \mathcal{F}_r$ is connected to the corresponding $\phi(A) \in \mathcal{K}_r$ and as a consequence the graph Γ' is connected precisely when the contraction of the graph Γ', given by contracting the pairs $\{A, \phi(A)\}$ to single vertices, is connected. Let G denote this new graph which is made up of $\binom{n}{r}$ vertices labelled by the r-subsets of X_n. In particular the vertices corresponding to the subsets A and B are connected if (but not only if) B may be obtained by adding the number 1 to one of the entries of A. Indeed, if $A_1 = \{i_1, \ldots, i_k, i_{k+1}, \ldots, i_r\}$, where $k < r$ and $i_{k+1} > i_k + 1$, and $A_2 = \{i_1, \ldots, i_k + 1, i_{k+1}, \ldots, i_r\}$, then A_2 is a transversal of the kernel $\phi(A_1)$ which means that vertex A_1 is adjacent to vertex A_2 in the graph G. As a consequence of this we can show that an arbitrary r-set $I = \{i_1, \ldots, i_r\}$ with $i_1 < \ldots < i_r$ is connected to $\{1, 2, \ldots, r\}$ since (with the symbol $\sim$ to be read as 'is connected to'):

$$\begin{array}{rcl} \{1,\ldots r-2, r-1, r\} \sim \{1,\ldots, r-2, r-1, r+1\} & \sim \ldots \sim & \{1, \ldots r-2, r-1, i_r\} \\ \sim \quad \{1, \ldots, r-2, r, i_r\} & \sim \ldots \sim & \{1, \ldots r-2, i_{r-1}, i_r\} \\ & \sim \ldots \sim & \{1, \ldots i_{r-2}, i_{r-1}, i_r\} \\ & \ddots & \vdots \\ & \sim & \{i_1, \ldots i_{r-2}, i_{r-1}, i_r\}. \end{array}$$

Therefore every vertex is connected to the vertex $\{1, \ldots, r\}$. We conclude that the graph G is connected and thus the graph Γ' is connected.

Secondly we have to check that Γ' has a symmetric distribution of idempotents with respect to the perfect matching ϕ (i.e. that for every $A \in \mathcal{F}_r$ we have $d(A) = d(\phi(A))$). On one hand if we fix the partition $K = \{[i_1, i_2 - 1], [i_2, i_3 - 1], \ldots, [i_r, i_1 - 1]\}$ then the number of images of size r that form a transversal of this partition is equal to the product $\prod_{j=1}^{r} |[i_j, i_{j+1} - 1]|$. On the other hand if we fix the image $I = \{i_1, \ldots, i_r\}$, with $i_1 \leq i_2 \leq \ldots \leq i_r$, and consider the partitions in $\phi(\mathcal{K}_r)$ that this image is a transversal of, we see that I is a transversal of $\{[j_1, j_2 - 1], [j_2, j_3 - 1], \ldots, [j_r, j_1 - 1]\}$, with $i_1 \in [j_1, j_2 - 1]$ say, if and only if $j_l \in [i_{l-1} + 1, i_l]$ (subscripts reduced mod r) for $1 \leq l \leq r$. We conclude that

$$d(\phi(A)) = \prod_{j=1}^{r} [i_j, i_{j+1} - 1] = \prod_{j=1}^{r} [i_j + 1, i_{j+1}] = d(A)$$

where all subscripts are reduced mod r. □

Theorem 3.35. *[59, Theorem 5] The semigroup $K(n,r) = \{\alpha \in T_n : |\operatorname{im}\alpha| \leq r\}$ satisfies*

$$\operatorname{idrank}(K(n,r)) = \operatorname{rank}(K(n,r)) = S(n,r)$$

for $1 < r < n$.

Proof. It follows from Lemma 3.34, Lemma 3.29 and Proposition 3.21 that P_r has an extremal idempotent generating set and then as a consequence of Lemma 3.30 so does $K(n,r)$. □

Finally note that the proof above is constructive and as a consequence we can write down a generating set of idempotents explicitly.

Corollary 3.36. *Let Y be the set of the maps*

$$\begin{pmatrix} 1 & 2 & \dots & i_1-1 & i_1 & \dots & i_2-1 & i_2 & \dots & i_r-1 & i_r & \dots & n-1 & n \\ i_r & i_r & \dots & i_r & i_1 & \dots & i_1 & i_2 & \dots & i_{r-1} & i_r & \dots & i_r & i_r \end{pmatrix}$$

where $1 \leq i_1 < i_2 < \dots < i_r \leq n$. For $i \in X_n$ let $C_\alpha(i) \subseteq X_n$ be the kernel class of α that contains i. Then with

$$Z = \{\alpha : \ker \alpha \neq \ker \beta \textit{ for all } \beta \in Y\} \cap \{\alpha : i\alpha = \min(C_\alpha(i)) \textit{ for } i \in X_n\}$$

the set of idempotents $Y \cup Z$ has cardinality $S(n,r)$ and it generates $K(n,r)$. □

The subsquare of J_r defined by the map ϕ was chosen carefully so that the distribution of idempotents was symmetric. If a subsquare were chosen at random then it may well fail to satisfy SHC.

Example 3.37. Let $n = 6$ and consider the following partitions of X_6 which all have weight 5.

$$\begin{array}{lllllll} k_1 & = & \{1,2\},\{3\},\{4\},\{5\},\{6\} & k_2 & = & \{1,3\},\{2\},\{4\},\{5\},\{6\} \\ k_3 & = & \{1,4\},\{2\},\{3\},\{5\},\{6\} & k_4 & = & \{2,3\},\{1\},\{4\},\{5\},\{6\} \\ k_5 & = & \{2,4\},\{1\},\{3\},\{5\},\{6\} & k_6 & = & \{3,4\},\{1\},\{2\},\{5\},\{6\} \end{array}$$

The set of all transversals is given by $X_6 \setminus \{i\}$ for $i = 1,2,3,4$. Therefore, in the graph $\Gamma(\mathcal{I},\mathcal{K})$, we have

$$|N(\{k_1,\dots,k_6\})| = 4 \leq 6 = |\{k_1,\dots,k_6\}|$$

and SHC is not satisfied. In fact, this subsquare is not even connected.

3.5.2 Ideals of the general linear semigroup

Now we will apply our results to give an analogous result to Theorem 3.35 but for the proper ideals of the monoid of endomorphisms of a finite vector space. In this case the $\mathcal{J}$-classes are square allowing us to apply Theorem 3.20 directly.

Let V be an n-dimensional vector space over a finite field F with $|F| = q$. Let $\mathrm{End}(V)$ denote the monoid of all linear transformations of V. Of course, $\mathrm{End}(V)$ can be represented concretely as the monoid of all $n \times n$ matrices with entries in the field F. The ideals of $\mathrm{End}(V)$ are given by

$$I(r,n,q) = \{A \in \mathrm{End}(V) : \dim(\mathrm{im}\, A) \leq r\} \text{ for } 1 \leq r \leq n$$

(see [21]). Denote the top $\mathcal{J}$-class of this subsemigroup, the $\mathcal{J}$-class consisting of all linear maps with $\dim(\mathrm{im}\, A) = r$, by $J(r,n,q)$ and the completely 0-simple semigroup $I(r,n,q)/I(r-1,n,q)$ by $PF(r,n,q)$. A summary of some known properties of $\mathrm{End}(V)$ is given below.

Proposition 3.38. *Let V be an n-dimensional vector space over the finite field F where $|F| = q$.*

(i) Green's relations are given by:

$$\begin{aligned} A\mathcal{R}B &\Leftrightarrow \mathrm{null}\, A = \mathrm{null}\, B; \\ A\mathcal{L}B &\Leftrightarrow \mathrm{im}\, A = \mathrm{im}\, B; \\ A\mathcal{J}B &\Leftrightarrow \dim(\mathrm{im}\, A) = \dim(\mathrm{im}\, B). \end{aligned}$$

(ii) Denote the $\mathcal{H}$-class consisting of elements A for which $\mathrm{im}\, A = I$ and $\mathrm{null}\, A = N$ by $H_{I,N}$. Then $H_{I,N}$ is a group if and only if $I \cap N = \{0\}$.

(iii) The number of non-zero $\mathcal{L}$-classes of $J(r,n,q)$ equals the number of non-zero $\mathcal{R}$-classes of $J(r,n,q)$ which equals the Gaussian coefficient

$$\begin{bmatrix} n \\ r \end{bmatrix}_q = \frac{(q^n-1)(q^{n-1}-1)\dots(q^{n-r+1}-1)}{(q^r-1)(q^{r-1}-1)\dots(q-1)}.$$

(iv) The number of idempotents in any non-zero $\mathcal{L}$-class of $J(r,n,q)$ equals the number of idempotents in any non-zero $\mathcal{R}$-class of $J(r,n,q)$ which equals $q^{r(n-r)}$.

(v) The semigroup $I(r,n,q)$ is idempotent generated and in particular is generated by the idempotents in its unique maximal $\mathcal{J}$-class $J(r,n,q)$.

Proof. (i) [21, Exercise 2.2.6] and [57, Exercise 2.16] (ii) [57, Exercise 2.19] (iii) This is just the formula for the number of r (and also $n-r$) dimensional subspaces of an n-dimensional vector space over a field F with size q (see [15, Chapter 9, Section 2]). (iv) A subspace of dimension r has $q^{r(n-r)}$ complements [11, Lemma 9.3.2.]. Also, a subspace of dimension $n-r$ has $q^{(n-r)(n-(n-r))} = q^{r(n-r)}$ complements. (v) This is a consequence of the main result of [29]. □

These facts along with the general results from §3.3 and §3.4 allow us to deduce the following.

Theorem 3.39. *Let V be an n-dimensional vector space over the finite field F where $|F| = q$. Then the semigroup $I(r,n,q) = \{A \in \mathrm{End}(V) : \dim(\mathrm{im}\, A) \leq r\}$ satisfies*

$$\mathrm{rank}(I(r,n,q)) = \mathrm{idrank}(I(r,n,q)) = \begin{bmatrix} n \\ r \end{bmatrix}_q$$

for $1 \leq r < n$.

Proof. By Proposition 3.38 the completely 0-simple semigroup $PF(r,n,q)$ has a $q^{r(n-r)}$-uniform distribution of idempotents which by Lemma 3.28 tells us that $PF(r,n,q)$ has an extremal idempotent generating set which in turn implies that $I(r,n,q)$ has an extremal idempotent generating set. □

As a byproduct of the above discussion we get the following result which characterises all generating sets of minimum cardinality of the semigroup $I(r,n,q)$.

Corollary 3.40. *A subset of $I(r,n,q)$ is a generating set of minimum cardinality for $I(r,n,q)$ if and only if it consists of $\begin{bmatrix} n \\ r \end{bmatrix}_q$ matrices all of whose images have dimension r no two of which have the same nullspace or the same image space.*

Proof. The completely 0-simple semigroup $PF(r,n,q)$ is square and has an extremal idempotent generating set. By Theorem 3.20 any sparse cover of $PF(r,n,q)$ generates $PF(r,n,q)$ which is equivalent to the statement in the corollary. □

We have seen that the graph $\Delta(PF(r,n,q))$ satisfies SHC for all $1 \leq r < n$. In fact, for the particular case when $r = n-1$, we can show that the graph $\Delta(PF(n-1,n,q))$ is Hamiltonian. The first step is to verify that $2q^{n-1} > (q^n - 1)/(q-1)$ (this follows from the fact that $q \geq 2$) which tells us that in the balanced bipartite graph $\Delta(PF(n-1,n,q))$, which has $2((q^n-1)/(q-1))$ vertices, every vertex has degree strictly greater than $((q^n-1)/(q-1))/2$. Then consider the following result, mentioned in Chapter 1, which gives a sufficient condition for a bipartite graph to have a Hamiltonian cycle.

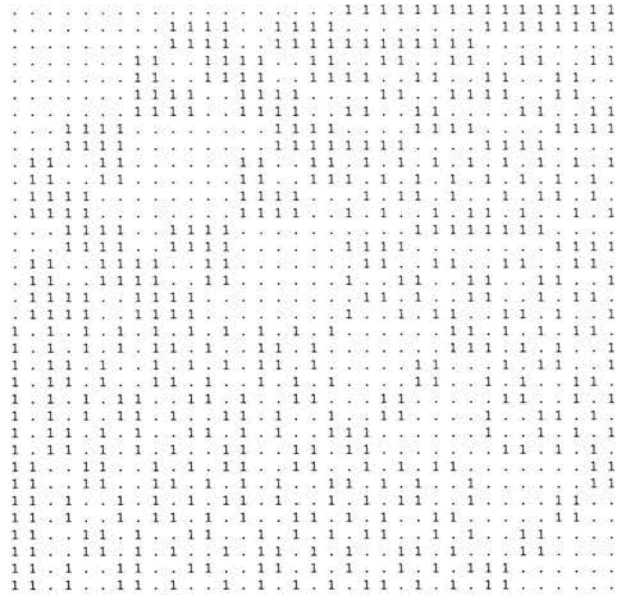


Figure 3.5: Egg-box picture of the unique maximal $\mathcal{J}$-class of the semigroup $I(2,4,2) \subseteq \mathrm{End}(\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2)$. It has 35 non-zero $\mathcal{R}$-classes, 35 non-zero $\mathcal{L}$-classes and 16 idempotents in every $\mathcal{R}$- and $\mathcal{L}$-class (represented by the 1s).

Theorem 3.41 (Moon and Moser, [75]). *If $G = (X, Y)$ is a bipartite graph with $|X| = |Y| = n$ such that for any non-adjacent pair of vertices $(x, y) \in V(X) \times V(Y)$ satisfies $d(x) + d(y) \geq n + 1$, then G has a Hamiltonian cycle.*

Proposition 3.42. *The graph $\Delta(PF(n-1, n, q))$ is Hamiltonian.*

Proof. Let $X \cup Y$ be the set of vertices of he balanced bipartite graph $\Delta(PF(n-1, n, q))$ with $|X| = |Y| = (q^n - 1)/(q - 1)$. Let $x \in X$ and $y \in Y$ be non-adjacent vertices in the graph. Then:

$$d(x) + d(y) = 2q^{n-1} > (q^n - 1)/(q - 1) = |X|.$$

It follows from Theorem 3.41 that $\Delta(PF(n-1, n, q))$ is Hamiltonian. □

The same approach cannot be used to prove the same thing for graphs associated with principal factors that are lower down in the semigroup.

Example 3.43. Let $V = \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$ and consider the semigroup PF_2. Then according to Proposition 3.38 the graph $\Delta(PF_2)$ is bipartite with $|I| = |\Lambda| = 35$ and each vertex has degree 16 (see Figure 3.5). In this example we cannot show that the graph is Hamiltonian by applying Moon and Moser's result since $16 + 16 = 32 < 35$.

This leaves open the following question.

Open Problem 4. Is $\Delta(PF(r,n,q))$ Hamiltonian for all $1 \le r < n$?

We have seen that it is true when $r = n-1$ and in [20] the author shows that it is true for some other special cases. The information given in Proposition 3.38 is not enough on its own to guarantee a Hamiltonian cycle. It is not true in general that every connected, k-regular bipartite graph necessarily has a Hamiltonian cycle. For example, in [28] an example is given of a 3-connected (a stronger condition than being connected), 3-regular bipartite graph that is non-Hamiltonian.

3.5.3 Transformation semigroups generated by mappings with prescribed image

In the introduction to this chapter the work of Levi and Seif was mentioned. In their work they say a partition of the set X_n has type $\tau = {d_1}^{\mu(d_1)}{d_2}^{\mu(d_2)} \dots {d_k}^{\mu(d_k)}$ if it has $\mu(d_i)$ classes of size d_i, where $d_1 > d_2 > \ldots > d_k$ and $n = \sum_{i=1}^{k} d_i\mu(d_i)$. The number of mappings with partition type τ is denoted by $\mathcal{N}(\tau)$. They consider the semigroups $S(\tau)$ generated by all transformations with partition type τ. In particular they prove the following.

Theorem 3.44. *[68, Theorem 1.10] Let n and r be positive integers with n greater than r, and let τ be a partition type of X_n of weight r. Then*

$$\operatorname{idrank}(S(\tau)) = \operatorname{rank}(S(\tau)) = \max(\mathcal{N}(\tau), \binom{n}{r}).$$

What happens if, rather than prescribing the kernels of the transformations, we prescribe the images? We begin with a result for completely 0-simple semigroups.

Theorem 3.45. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be an idempotent generated completely 0-simple semigroup with $\operatorname{rank}(S) = \operatorname{idrank}(S)$ and $|I| \le |\Lambda|$. For $\mathcal{A} \subseteq I$ define:*

$$S(\mathcal{A}) = \{(i, g, \lambda) : i \in \mathcal{A},\ g \in G,\ \lambda \in \Lambda\} \cup \{0\}$$

a, not necessarily regular, Rees matrix subsemigroup of S. If $\operatorname{rank}(G) \le 2$ then

$$\operatorname{rank}(S(\mathcal{A})) = \max(|I|, |\Lambda|) = |\Lambda|.$$

Proof. The generalised Rees matrix semigroup $S(\mathcal{A})$, in general, will not be regular. It will, however, have the property that for every $i \in \mathcal{A}$ there exists at least one $\lambda \in \Lambda$ such that $p_{\lambda i} \neq 0$. In this sense it is semi-regular. That is, the matrix P has the form

$$\begin{array}{cc} & \mathcal{A} \\ \begin{array}{c} \Lambda_1 \\ \\ \Lambda_2 \end{array} & \left(\begin{array}{c} Q \\ \hline \\ \end{array}\right) \end{array}$$

where Q is regular and all other entries equal zero. Let $T = \mathcal{M}^0[G; \mathcal{A}, \Lambda_1; Q]$. Since S has an extremal idempotent generating set, by Theorem 3.27, for every subset Y of $\mathcal{A}$ we have $|N(Y)| > |Y|$. In particular it follows that Q is not the identity matrix and moreover that Q has strictly more rows than it has columns (as suggested in the picture above). Since $\text{rank}(G) \leq 2$ it follows that in the formula for $\text{rank}(T)$ (given by Theorem 2.51) we have $r_{\min} \leq 2$. Also, the matrix Q has at most $|\mathcal{A}| < |\Lambda_1|$ connected components. It follows that:

$$r_{\min} + k - 1 < 2 + |\Lambda_1| - 1 = |\Lambda_1| + 1$$

and therefore by Theorem 2.68

$$\begin{aligned} \text{rank}(S(\mathcal{A})) &= \text{rank}(T) + |\Lambda_2| = \max(|\mathcal{A}|, |\Lambda_1|, r_{\min} + k - 1) + |\Lambda_2| \\ &= |\Lambda_1| + |\Lambda_2| = \Lambda \end{aligned}$$

as required. □

There is an obvious dual result to the one above for the case $|\Lambda| \geq |I|$ and $\mathcal{A} \subseteq \Lambda$. Since the finite symmetric and general linear groups are both 2-generated (see [94]) the following corollaries are obtained.

Theorem 3.46. *Let $\mathcal{A}$ be a set of r-subsets of $\{1, \ldots, n\}$ where $1 < r < n$. Let $S(\mathcal{A})$ be the semigroup generated by the set of all mappings α with $\text{im}\,\alpha \in \mathcal{A}$. Then:*

$$\text{rank}(S(\mathcal{A})) = S(n, r).$$

Proof. Let $S = K(n,r)/K(n, r-1)$ and let $I = \{\alpha \in S(\mathcal{A}) : |\text{im}\,\alpha| < r\}$. Then $S(\mathcal{A})/I$ is isomorphic to a subsemigroup of S of the form described in

Theorem 3.45. By Theorem 3.35 the semigroup S has an extremal idempotent generating set. Moreover, the symmetric group is 2-generated. Thus all the conditions of Theorem 3.45 are satisfied and the result follows. □

It is a little surprising that the rank of $S(\mathcal{A})$ only depends on n and r and is independent of the particular set $\mathcal{A}$ of r-subsets chosen. The exact analogue of the above results holds for the ideals of the general linear semigroup.

Theorem 3.47. *Let V be a finite vector space with dimension n over the finite field F with $|F| = q$. Let $\mathcal{F}$ be a family of subspaces of V each with dimension r for some $1 < r < n$. Let $S(\mathcal{F})$ be the subsemigroup of $\operatorname{End}(V)$ generated by the set of all linear transformations α with $\operatorname{im}\alpha \in \mathcal{F}$. Then*

$$\operatorname{rank}(S(\mathcal{F})) = \begin{bmatrix} n \\ r \end{bmatrix}_q .$$

Proof. Since $I(r, n, q)$ has an extremal idempotent generating set and the general linear group is two generated the conditions are satisfied for applying Theorem 3.45.

□

3.6 Nilpotent rank

In [87] it was suggested by Schwarz that the role of *nilpotents* in semigroups should be investigated further. Recall that given a semigroup S with zero $0 \in S$ we say $s \in S$ is nilpotent if $s^n = 0$ for some $n \in \mathbb{N}$. Let $N(S)$ denote the set of nilpotents of a semigroup S.

In [91] it was shown that the semigroup SP_n of strictly partial transformations on X_n is nilpotent generated if n is even, and if n is odd then $\langle N(SP_n)\rangle = SP_n \setminus W_{n-1}$ where W_{n-1} consists of all elements J_{n-1} whose *completions* are odd permutations. Similarly, in [38] Gomes and Howie showed that for n even, the semigroup of proper subpermutations $SI_n \leq I_n$ is nilpotent generated, and for n odd the nilpotents generate $SI_n \setminus W_{n-1}$. One major difference between analysing the subsemigroup generated by the nilpotent elements, as opposed to that generated by the idempotents, is that if e is an idempotent in the principal factor J_e^* then it is an idempotent in S. The same is not, however, true of nilpotent elements.

Example 3.48. Let $\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \end{pmatrix} \in P_3$ the partial transformation semigroup on X_3. The zero of P_n is $z = \begin{pmatrix} 1 & 2 & 3 \\ - & - & - \end{pmatrix}$. Clearly no power of α is equal to z so α is not a nilpotent element in P_3. On the other hand, in the principal factor $J_\alpha{}^*$ we have $\alpha^2 = 0$.

By analogy with the rank and idempotent rank we define the nilpotent rank as

$$\text{nilrank}(S) = \min\{|B| : B \subseteq N(S), \langle B \rangle = S\}.$$

Because of examples like the one immediately above, in most cases, determining the nilpotent rank of a semigroup does not reduce to determining the nilpotent rank of its principal factors. In the same way as for idempotent rank, the nilpotent rank of S and the rank of S are not going to be the same in general.

Example 3.49. The rectangular 0-band with structure matrix

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

has rank 3 and nilpotent rank 4.

One of the earliest results on nilpotent rank was given by Gomes and Howie in [38] where they show that $SI_n \leq I_n$ satisfies:

$$\text{rank}(SI_n) = \text{nilrank}(SI_n) = n + 1$$

for n even and $n \geq 4$. Some other nilpotent generated semigroups and their nilpotent ranks are given in Table 3.1. In this section we will consider the question of nilpotent rank for arbitrary finite completely 0-simple semigroups.

As we did for idempotent rank we want to relate the question of nilpotent rank of S to the study of the nilpotent rank of its natural rectangular 0-band homomorphic image. It is not quite as simple as the idempotent rank case here though. We make use of Proposition 2.49, from the previous chapter, which tells us how we can build generating sets for S around the coordinates of generating sets of $T = S\natural$.

Theorem 3.50. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a nilpotent generated completely 0-simple semigroup. Let T be the rectangular 0-band homomorphic image of S. Then*

$$\text{nilrank}(S) = \max(\text{rank}(S), \text{nilrank}(T)).$$

Semigroup	Description	Rank	Idrank	Nilrank	Reference
Sing_n	$\alpha \in T_n : \lvert \mathrm{im}\,\alpha \rvert \le n-1$	$\frac{n(n-1)}{2}$	$\frac{n(n-1)}{2}$	–	[37]
$\mathrm{Sing}(V)$	$\alpha \in \mathrm{End}(V) : \lvert \dim(\mathrm{im}\,\alpha) \rvert \le n-1$	$\frac{q^n-1}{q-1}$	$\frac{q^n-1}{q-1}$	$\frac{q^n-1}{q-1}$	[26] and Theorem 3.55
$K(n,r)$	$\alpha \in T_n : \lvert \mathrm{im}\,\alpha \rvert \le r$	$S(n,r)$	$S(n,r)$	–	[59]
$I(r,n,q)$	$\alpha \in \mathrm{End}(V) : \lvert \dim(\mathrm{im}\,\alpha) \rvert \le r$	$\begin{bmatrix} n \\ r \end{bmatrix}_q$	$\begin{bmatrix} n \\ r \end{bmatrix}_q$	$\begin{bmatrix} n \\ r \end{bmatrix}_q$	Theorems 3.39 and 3.55
$KP(n,r)$	$\alpha \in P_n : \lvert \mathrm{im}\,\alpha \rvert \le r$	$S(n+1,r+1)$	$S(n+1,r+1)$	–	[34]
O_n	$\alpha \in \mathrm{Sing}_n : (\forall x,y \in X_n)$ $x \le y \Rightarrow x\alpha \le y\alpha$	n	$2n-2$	–	[38]
PO_n	$O_n \cup \{\alpha : \mathrm{dom}(\alpha) \subsetneq X_n,$ $(\forall x,y \in \mathrm{dom}(\alpha))\ x \le y \Rightarrow x\alpha \le y\alpha\}$	$2n-1$	$3n-2.$	–	[38]
SPO_n	$PO_n \setminus O_n$	$2n-2$	??	??	[38]
OP_r	principal factor of O_n	$\binom{n}{r}$	$\binom{n}{r}$ (for $2 \le r \le n-2$)	$\binom{n}{r}$	[96]
$U(n,r)$	$\alpha \in SI_n : \lvert \mathrm{im}\,\alpha \rvert \le r$	$\binom{n}{r}+1$	–	$\binom{n}{r}+1$	[36]
$V(n,r)$	$\alpha \in SP_n : \lvert \mathrm{im}\,\alpha \rvert \le r$	$(r+1)S(n,r+1)$	??	$(r+1)S(n,r+1)$	[36]
P_r	$K(n,r)/K(n,r-1)$	$S(n,r)$	$S(n,r)$	$S(n,r)$	[95]
$S(\tau)$	$\langle \alpha \in T_n : \alpha \text{ has type } \tau \rangle$	$\max\{\mathcal{N}(\tau), \binom{n}{r}\}$	$\max\{\mathcal{N}(\tau), \binom{n}{r}\}$	–	[67]
PF_r	principal factor of P_n	$S(n+1,r+1)$	$S(n+1,r+1)$	$S(n+1,r+1)$	[97]
OF_r	principal factor of PO_n	$\sum_{k=r}^{n} \binom{n}{k}\binom{k-1}{r-1}$	$\sum_{k=r}^{n} \binom{n}{k}\binom{k-1}{r-1}$	$\sum_{k=r}^{n} \binom{n}{k}\binom{k-1}{r-1}$	[97]

Table 3.1: Known results concerning ranks of semigroups. We write – when the question does not apply, for example Sing_n does not have a zero element and so it doesn't make sense to ask for the nilpotent rank. We write ?? if the answer is unknown.

Proof. ($\geq$) Let A be a nilpotent basis for S. Then $|A| \geq \text{rank}(S)$ and since $A\natural \subseteq T$ is a nilpotent generating set of T it follows that $|A| \geq \text{nilrank}(T)$. ($\leq$) By Proposition 2.49 there exists a generating set A' of S such that $A'\natural \subseteq A$ where A is a nilpotent generating set for T with $|A| = \text{nilrank}(T)$ and

$$|A'| = \max(\text{rank}(S), |A|) = \max(\text{rank}(S), \text{nilrank}(T)).$$

Since $A \subseteq N(T)$ it follows that $A' \subseteq N(S)$. □

As an immediate consequence we have

Corollary 3.51. *If $S = \mathcal{M}^0[G; I, \Lambda; P]$ is both idempotent and nilpotent generated and $T = S\natural$ then*

$$\text{nilrank}(S) = \text{nilrank}(T).$$

Proof. Since S is idempotent generated it follows from Theorem 2.13 that $\text{rank}(S) = \max(|I|, |\Lambda|)$. Therefore by Theorem 3.50

$$\text{nilrank}(S) = \max(|I|, |\Lambda|, \text{nilrank}(T)) = \text{nilrank}(T).$$

□

We may now determine the nilpotent rank of the principal factors of some of the examples in Table 3.1.

Corollary 3.52. *[95, Theorem 5] The principal factors $P_r = K(n,r)/K(n,r-1)$ all satisfy:*

$$\text{rank}(P_r) = \text{idrank}(P_r) = \text{nilrank}(P_r) = S(n,r).$$

Proof. We saw the proof of the rank and idempotent rank results in Theorem 3.35. Take the subsquare of the unique maximal $\mathcal{J}$-class of $K(n,r)$ described in Lemma 3.34. The dual of the bipartite graph $\Gamma' = \mathcal{F}_r \cup \phi(\mathcal{F}_r)$ also has a uniform distribution of edges. By the same arguments we can find a perfect matching which can be used to pick $\binom{n}{r}$ non-group elements which form a sparse cover of the subsquare. This sparse cover generates a transversal of the $\mathcal{H}$-classes of $\mathcal{F}_r \cup \phi(\mathcal{F}_r)$ by Theorem 3.20. Taking these generators along with arbitrary non-group $\mathcal{H}$-class representatives, one from every remaining $\mathcal{R}$-class or P_r, gives a generating set with $S(n,r)$ elements all of which belong to non-group $\mathcal{H}$-classes of J_r and so are nilpotents in $J_r{}^*$. □

Note that the question of the nilpotent rank of $K(n,r)$ itself does not arise since $K(n,r)$ does not have a zero. A related question would be to consider generating with elements α such that for some natural number k the element α^k is a constant.

Returning to the Brandt semigroup we recover the following result:

Proposition 3.53. *[97, Theorem 4.1] Let $B = B(G, \{1, \ldots, n\})$ be a Brandt semigroup, where G is a group of rank r ($r \geq 1$). Then*

$$\text{nilrank}(B) = \text{rank}(B) = n + r - 1.$$

Proof. The natural rectangular 0-band homomorphic image T is the aperiodic Brandt semigroup which satisfies $\text{nilrank}(T) = n$ (in fact all bases of this zero rectangular band are made up of nilpotents). By Theorem 3.50 $\text{nilrank}(B) = \max(\text{nilrank}(T), \text{rank}(B)) = \max(n, n + r - 1) = n + r - 1$. □

We now determine the nilpotent rank of the ideals of the general linear semigroup. We must first show that they are nilpotent generated. We rely on the fact that it is idempotent generated in order to prove this. The key to proving the result is the observation that every non-group $\mathcal{H}$-class of $\text{End}(V)$ has at least one nilpotent element. This, along with the fact that it has an extremal idempotent generating set, allows the problem to be reduced to a corresponding question for completely 0-simple semigroups.

Lemma 3.54. *Let I be a subspace of V with $\dim V = n$ and $\dim I = r$. Let N be a subspace of V with dimension $n - r$. If $I \cap N \neq \{0\}$ then the $\mathcal{H}$-class $H_{I,N}$ contains at least one nilpotent element.*

Proof. Let $W = I \cap N$ and, say, $\dim W = k$ where $1 \leq k \leq \min(r, n - r)$. Let $B_W = \{w_1, \ldots, w_k\}$ be a basis for W. Define $B_N = \{\nu_1, \ldots, \nu_{n-r-k}\}$ and $B_I = \{\iota_1, \ldots, \iota_{r-k}\}$ so that $B_W \cup B_N$ is a basis for N and $B_W \cup B_I$ is a basis for I. Note that $B_W \cup B_N \cup B_I$ is an independent subset of V with $|B_W \cup B_N \cup B_I| = n - k$. This is because it has $n - k$ elements and generates $I + N$. But $\dim(I + N) = \dim(I) + \dim(N) - \dim(I \cap N) = n - k$ (see [77, Theorem 2.15]) so the set must be independent. Now extend $B_W \cup B_N \cup B_I$ by $C = \{c_1, \ldots, c_k\}$ to a basis $B_W \cup B_N \cup B_I \cup C$ of V. Define $\alpha \in \text{End}(V)$ to be the unique linear transformation extending the map

$$\begin{pmatrix} \iota_1 & \iota_2 & \ldots & \iota_{r-k-1} & \iota_{r-k} & c_1 & \ldots & c_{k-1} & c_k & w_1 & \ldots & w_k & \nu_1 & \ldots & \nu_{n-r-k} \\ \iota_2 & \iota_3 & \ldots & \iota_{r-k} & w_1 & w_2 & \ldots & w_k & \iota_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

Then $\text{im}\,\alpha = \langle B_W \cup B_I \rangle = I$ and $\text{null}\,\alpha = \langle B_W \cup B_N \rangle = N$ so $\alpha \in H_{I,N}$. Also $\alpha^{r-k+1} = 0$, since it sends every basis element to 0, and hence is nilpotent. □

Theorem 3.55. *Let V be an n-dimensional vector space over the finite field F where $|F| = q$. Then the semigroup $I(r, n, q) = \{A \in \mathrm{End}(V) : \dim(\mathrm{im}\, A) \leq r\}$ satisfies*

$$\mathrm{nilrank}(I(r, n, q)) = \begin{bmatrix} n \\ r \end{bmatrix}_q$$

for $1 \leq r < n$.

Proof. Let $S = I(r, n, q)/I(r-1, n, q)$ and let $M = \begin{bmatrix} n \\ r \end{bmatrix}_q$. By Theorem 3.20 any sparse cover of S is a generating set for S. The complement of the graph $\Delta(S)$ is a regular bipartite graph and so has a perfect matching. Let $H_1, \ldots, H_M$ be the non-group $\mathcal{H}$-classes of $J_r \subseteq I(r, n, q)$ corresponding to the edges in this perfect matching. By Lemma 3.54 each H_j contains at least one nilpotent linear transformation. Let $\alpha_j \in H_j$ be a nilpotent linear transformation for $1 \leq j \leq M$. The subset $\{\alpha_j : 1 \leq j \leq M\}$ of $N(I(r, n, q))$ is a sparse cover of $J_r{}^*$ and so generates $J_r{}^*$, by Theorem 3.20, and hence generates the whole of $I(r, n, q)$ by Proposition 3.38. □

3.7 Counting generating sets

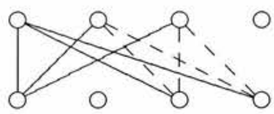
One of the surprising things highlighted by Theorem 3.20 is the connection between the idempotent rank of S and the problem of counting the number of generating sets with minimum cardinality. These connections will be investigated further in this section.

Given a rectangular 0-band S with structure matrix P we use $S[i][\lambda]$ to denote the (not necessarily regular) rectangular 0-band with structure matrix $P[\lambda][i]$.

Definition 3.56. Let S be a rectangular 0-band with structure matrix P and let $s = (j, \mu) \in S$. Define $T(s, S)$ to be the rectangular 0-band with the same dimensions as S and with structure matrix $Q = (q_{\lambda i})$ defined by:

$$q_{\lambda i} = \begin{cases} 1 & \text{if } p_{\lambda i} = 1 \text{ or } p_{\mu,i} = p_{\lambda,j} = 1 \\ 0 & \text{otherwise.} \end{cases}$$

We call $T(s, S)$ the *full completion* of S with respect to s. The graph $\Delta(Q)$ is constructed from the graph $\Delta(P)$ by adding in various edges. The following diagram illustrates this, the dotted edges in the diagram are those that are being added in.



We have the following basic properties:

Proposition 3.57. *Let S be a rectangular 0-band and let $s = (j, \mu) \in S$ be an element such that there are idempotents e and f with $e \neq s$, $f \neq s$, $s\mathcal{R}e$ and $s\mathcal{L}f$. Then:*

(i) $T(s, S)[j][\mu]$, $T(s, S)[0][\mu]$ and $T(s, S)[j][0]$ are all regular;

(ii) for any subset A of S with $s \in A$ we have $\langle A \rangle_S = \langle A \rangle_{T(s,S)}$.

Proof. (i) We start by proving that $T(s, S)[0][\mu]$ is regular. This is equivalent to showing that the bipartite graph $\Delta(T(s, S)[0][\mu])$ has no isolated vertices. This graph has vertex set $I \cup (\Lambda \setminus \{\mu\})$. None of the vertices in $\Lambda \setminus \{\mu\}$ are isolated in $\Delta(T(s, S)[0][\mu])$ since none of them are isolated in $\Delta(S)$. None of the vertices in $I \setminus N(\mu)$ (the neighbourhood of μ) are isolated for the same reason. By assumption, $N(j) \setminus \{\mu\}$ is non-empty and every vertex of $N(j) \setminus \{\mu\}$ is connected to every vertex of $N(\mu)$ in $\Delta(T(s, S)[0][\mu])$. Therefore, none of the vertices of $\Delta(T(s, S)[0][\mu])$ are isolated and so $T(s, S)[0][\mu]$ is regular. The fact that $T(s, S)[j][0]$ is regular follows from a dual argument. The semigroup $T(s, S)[j][\mu]$ is regular since $T(s, S)[0][\mu]$ and $T(s, S)[j][0]$ both are. (ii) Follows from the definitions combined with Lemma 2.44. □

In Theorem 2.10 is was shown that for every finite $n \times n$ rectangular 0-band S we can always find at least one set A with $|A| = n$ such that $\langle A \rangle = S$. In fact we can say much more than this.

Theorem 3.58. *Let S be an $n \times n$ rectangular 0-band. Let:*

$$\mathcal{B}(S) = \{A \subseteq S : |A| = n \ \& \ \langle A \rangle = S\}.$$

Then $|\mathcal{B}(S)| \geq (n-1)!$. Moreover, there exist examples where this lower bound is attained.

Proof. We prove the result by induction on the dimension of S. When $n = 1$ the result holds trivially since in this case $|\mathcal{B}(S)| = 1 = 0!$. Now let S have dimension

k and suppose that the result holds for all smaller examples. There are two cases to consider:

Case 1: *S is equivalent to the $k \times k$ aperiodic Brandt semigroup.* In this case the result follows from Lemma 2.23.

Case 2: *S is not equivalent to the $k \times k$ aperiodic Brandt semigroup.* In this case by Lemma 2.7 there exist $i, \lambda \in \{1 \ldots, k\}$ such that $S[0][\lambda]$, $S[i][0]$ and $S[i][\lambda]$ are all regular. Let (i, μ) be an idempotent in S with $\mu \neq \lambda$. Such an idempotent must exist since $S[0][\lambda]$ is regular. Now define:

$$X = \{(i, \xi) \in S : \xi \neq \mu\}.$$

Claim. *For every $x \in X$ there are at least $(n-2)!$ sets $B \in \mathcal{B}(S)$ such that $x \in B$.*

Proof. Let $x \in X$. If $x = s = (i, \lambda)$ then the $(k-1) \times (k-1)$ rectangular 0-band $S[i][\lambda]$ is regular by definition, so we can apply induction constructing the set $\{C_1, \ldots, C_{(k-2)!}\}$ of generating sets for $S[i][\lambda]$. Each of the sets $\{x\} \cup C_y$ for $y \in \{1, \ldots, (k-2)!\}$ is a generating set for S and they are all distinct.

On the other hand, if $x = (i, \xi) \neq (i, \lambda)$ then $S[i][0]$ is regular, by definition, but $S[0][\xi]$ and $S[i][\xi]$ may not be regular. Now construct $S(x, S)$ noting that, by Proposition 3.57, all of $S(x, S)[i][0]$, $S(x, S)[0][\xi]$ and $S(x, S)[i][\xi]$ are regular. By induction the rectangular 0-band $S(x, S)[i][\xi]$ has at least $(n-2)!$ distinct bases $\{C_1, \ldots, C_{(k-2)!}\}$. Each of the sets $\{x\} \cup C_y$ for $y \in \{1, \ldots, (k-2)!\}$ is a generating set for $S(x, S)$ and they are all distinct. But then, by Proposition 3.57, we have

$$\langle \{x\} \cup C_y \rangle_{S(x,S)} = \langle \{x\} \cup C_y \rangle_S$$

and so $\{x\} \cup C_y$ are all bases of S. □

Combining all of these generating sets together we get $(n-1)(n-2)! = (n-1)!$ distinct bases for S. □

Example 3.59. Let S be an $n \times n$ rectangular band. Clearly

$$A_\alpha = \{(i, i\alpha) : 1 \leq i \leq n\}$$

generates S for all $\alpha \in S_n$ and so $|\mathcal{B}(S)| = n!$ in this case.

Example 3.60. Let S be the rectangular $n \times n$ rectangular 0-band with structure matrix

$$P = \begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Since $(1, 1) \notin \langle S \setminus \{(1, 1)\} \rangle$ this element must be included in every generating set of S. Then the set $\{(1, 1), (2, 2\alpha), (3, 3\alpha), \dots (n, n\alpha)\}$ generates S for every $\alpha \in S(\{2, 3, \dots, n\})$. Therefore $|\mathcal{B}(S)| = |S_{n-1}| = (n - 1)!$ in this case.

This leads to the following result.

Proposition 3.61. *Let S be an $m \times n$ rectangular 0-band where $m \leq n$. Then*

$$(m-1)!\, m^{(n-m)} \leq |\mathcal{B}(S)| \leq m!\, S(n, m).$$

Proof. By Lemma 2.7 the semigroup S has an $m \times m$ regular sub-rectangular 0-band which by Theorem 3.58 has at least $(m - 1)!$ distinct bases. For each of these bases we can choose arbitrarily from the remaining columns, giving a basis for S. We conclude that $|\mathcal{B}(S)| \geq (m - 1)! m^{n-m}$. The upper bound is just the number of onto maps from an n-set to an m-set. □

This gives a corresponding lower bound for the number of bases of an arbitrary finite completely 0-simple semigroup.

Proposition 3.62. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a finite completely 0-simple semigroup where $|I| \leq |\Lambda|$. Then*

$$|\mathcal{B}(S)| \geq (|I| - 1)!\, |I|^{(|\Lambda| - |I|)}.$$

Proof. Let T be the natural rectangular 0-band homomorphic image of S. By Proposition 3.61 the semigroup T has at least $(|I| - 1)!\, |I|^{(|\Lambda| - |I|)}$ distinct bases. By Proposition 2.49 corresponding to every basis of T there is at least one basis of S with the same coordinates. □

In Theorem 3.20 we saw that square rectangular 0-bands have extremal idempotent generating sets if and only if every sparse cover is a generating set. This is not true for arbitrary, non-square, examples.

Example 3.63. Let S be the rectangular 0-band with structure matrix

$$P = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Then

$$A = \{(4,1),(3,2),(2,3),(1,4),(1,5),(1,6),(1,7),(1,8)\}$$

is an idempotent basis for S. However, the set

$$B = \{(4,1),(4,2),(4,3),(4,4),(4,5),(3,6),(2,7),(1,8)\}$$

is a sparse cover of S and does not generate S.

This leaves us the problem of describing all non-square completely 0-simple semigroups that are generated by all their sparse covers. Can a completely 0-simple semigroup that is not idempotent generated have this property? We seek the answer to this question below.

Lemma 3.64. *Let S be a finite completely 0-simple semigroup. If S is generated by all of its sparse covers then S is idempotent generated.*

Proof. Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a completely 0-simple semigroup that is not idempotent generated. Suppose that S is in Graham normal form. Then either S is not connected, or S is connected and the non-zero entries of P do not generate G.

Case 1: *S is not connected.* If S is not connected then it is easy to find a sparse cover that does not generate S. Indeed, since S is not connected the matrix P can be assumed without loss of generality to have the form:

$$\begin{pmatrix} A & O_1 \\ O_2 & B \end{pmatrix}$$

where the entries in O_1 and O_2 are all equal to zero. Also assuming without loss of generality that $I = \{1, \ldots, n\}, \Lambda = \{1, \ldots, m\}$ and $n \leq m$ we let

$$X = \{(1, 1_G, 1), (2, 1_G, 2), \ldots, (n, 1_G, n), (n, 1_G, n+1), (n, 1_G, n+2), \ldots, (n, 1_G, m)\}.$$

Then X intersects A and B, and either O_1, or O_2, but not both. If X intersects O_1 then $\langle X \rangle$ does not intersect O_2 and visa-versa. Therefore X is a sparse cover of S that does not generate S.

Case 2: *S is connected.* In this case let

$$A = \{(i_1, 1_G, \lambda_1), \ldots, (i_r, 1_G, \lambda_r)\}$$

be a sparse cover of S such that the middle components are all equal to the identity of the group G. Suppose without loss of generality that $p_{11} = 1$. Since

$$\langle A \rangle \cap H_{11} \subseteq \{(1, h, 1) : h \in H\} \subsetneq 1 \times G \times 1$$

where H is the subgroup generated by the non-zero entries in the matrix P, it follows that A does not generate S. □

Theorem 3.65. *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a finite completely 0-simple semigroup. Then every sparse cover of S is a generating set for S if and only if S is idempotent generated and every square extension of S satisfies SHC.*

Proof. ($\Leftarrow$) Let $\Sigma \subseteq S$ be a sparse cover of S. Say $\Sigma = (A_1, 1) \cup (A_2, 2) \cup \ldots \cup (A_n, n)$ where, without loss of generality, we suppose that $i \in A_i$ for all i. Let T_Σ be the square extension of S constructed using the family $\mathcal{F} = \{A_1, \ldots, A_n\}$. By assumption T_Σ satisfies SHC. Fix bijections $\phi_i : A_i \to A_i$ for all i and use these maps to construct a bijection $\phi : \{1, \ldots, m\} \to \{1, \ldots, m\}$. Since T_Σ satisfies SHC and $B = \{(i, \phi(i)) : i \in \{1 \ldots, m\}\}$ is a sparse cover of T_Σ it follows from Theorem 3.20 that $\langle B \rangle = T_\Sigma$. Let $\psi : T \to S$ be the onto homomorphism defined by $(a, b)\psi = (a, \overline{f}(b))$ (i.e. the retraction mapping associated with the square extension $\mathcal{F}$). Then $B\psi = \Sigma$ and $\langle B \rangle = S$ which implies that $\langle B\psi \rangle = T$ since ψ is a homomorphism. Therefore $\langle \Sigma \rangle = T$ which completes the proof since Σ was an arbitrary sparse cover of S.

($\Rightarrow$) Suppose that every sparse cover of S is a generating set for S. It follows from Lemma 3.64 that S is idempotent generated. Since S is idempotent generated, it follows from Lemma 3.11 that it is sufficient to prove the result under the assumption that S is a rectangular 0-band. Therefore let S be a $m \times n (m \geq n)$ rectangular 0-band such that every sparse cover of S is a generating set.

Claim. *Every non-zero $\mathcal{L}$-class of S has at least $m - n + 1$ idempotents.*

Proof. Suppose otherwise and seek a contradiction. Say column 1 has x idempotents where $x < m-n+1$. Without loss of generality say $\{(1,1),(2,1),\ldots,(x,1)\}$ are all the idempotents in this $\mathcal{L}$-class. Let

$$Y = \{(1,1),(2,1),\ldots,(x,1),(x+1,2),(x+2,3),\ldots,$$

$$(x+(n-1),n),(x+n,n),(x+n+1,n),(x+n+2,n),\ldots,(m,n)\},$$

which is is a sparse cover of S. However, $\langle Y \rangle \neq S$ since in particular $(1,n) \notin \langle Y \rangle$. This contradicts the assumption that every sparse cover is a generating set, completing the proof of the claim. □

Now consider an arbitrary square extension T of S and an arbitrary sparse cover C of T thinking of S as a subsemigroup of T. Using the claim we know that every $\mathcal{L}$-class has strictly more that $m - n$ elements. It follows that C generates a sparse cover of S inside T. By assumption this sparse cover generates S. Now using the elements in S and the elements in $C \setminus S$ we can, by regularity of S, generate the whole of T. Since the original sparse cover was arbitrary we conclude that every sparse cover of T generates T which, by Theorem 3.20, implies that T satisfies SHC. □

We saw in Corollary 3.40 that every sparse cover of $I(r,n,q)/I(r-1,n,q)$ is a generating set. In general, the same is not true for the semigroup $K(n,r)/K(n,r-1)$.

Example 3.66. The number of idempotents in every $\mathcal{R}$-class of $J_r \subseteq T_n$ is r^{n-r}. Whenever $S(n,r) - \binom{n}{r} > r^{n-r}$ we can find a square extension of $J_r{}^*$ that does not satisfy SHC. Consider $K(5,3)$ for example. In this example every $\mathcal{R}$-class of J_3 has 9 idempotents, $S(5,3) = 25$ and $\binom{5}{3} = 10$. If we extend $J_3{}^*$ to a square by repeatedly adjoining copies of a single $\mathcal{R}$-class then, since $9 < 15$, the resulting square extension will not satisfy SHC. We conclude that not every sparse cover of $J_3 \subseteq T_5$ generates $J_3{}^*$.

Using Corollary 3.40 we can count the number of bases of the semigroup $I(r,n,q)$.

Proposition 3.67. *Let F be a finite field with size q and let S be the general linear semigroup $GLS(n,q)$. Let $I(r,n,q)$ be a two-sided ideal of $GLS(n,q)$. Then the number of generating sets of $I(r,n,q)$ with minimum cardinality is given by*

$$|\mathcal{B}(I(r,n,q))| = \begin{bmatrix} n \\ r \end{bmatrix}_q ! \, [(q^r-1)(q^r-q)\dots(q^r-q^{r-1})]^{\begin{bmatrix} n \\ r \end{bmatrix}_q}.$$

On the other hand, for $K(n,r)$ we still do not have an exact formula for the number of bases. Using Lemma 3.34 we may obtain a lower bound.

Proposition 3.68. *Let $n, r \in \mathbb{N}$ with $2 \le r \le n-1$. Then a lower bound for the number of generating sets of $K(n,r)$ with minimum cardinality is given by*

$$|\mathcal{B}(K(n,r))| \ge \binom{n}{r}! \, (r!)^{S(n,r)} \left(S(n,r) - \binom{n}{r} \right)^{\binom{n}{r}}.$$

Proof. Consider the subsemigroup of $P_r = K(n,r)/K(n.r-1)$ which is the union of the $\mathcal{R}$-classes indexed by the partitions defined in Definition 3.31. Let T be the natural rectangular 0-band homomorphic image of this semigroup. Then T is a square rectangular 0-band that satisfies SHC. Therefore T has $\binom{n}{r}!$ distinct generating sets. Let S be the natural rectangular 0-band homomorphic image of $K(n,r)/K(n,r-1)$. The semigroup T is a subsemigroup of S and every generating set of T may be extended to a generating set for S in $(S(n,r)-\binom{n}{r})^{\binom{n}{r}}$ different ways. Let $\mathcal{X}$ be the family of all such generating sets of S. Since P_r is idempotent generated, for every $X \in \mathcal{X}$ the coordinates X may be extended to a generating set of triples for P_r in $r!^{|X|} = (r!)^{S(n,r)}$ different ways. □

Thus, many of the sparse covers do generate $J_r{}^*$. On the other hand, not all of the sparse covers do. This leaves the following problem.

Open Problem 5. Describe all generating sets of $K(n,r)$ with minimal cardinality, and then count them.

Chapter 4

Free G-sets and trivial independence algebras

4.1 Introduction

For any mathematical object M the set $\mathrm{End}(M)$ of all endomorphisms of M is closed under composition and forms a monoid. For example, the semigroup of order preserving mappings O_n, introduced in Section 2.1, may be viewed as the endomorphism monoid of a linearly ordered set with n elements. The general linear semigroup $\mathrm{GLS}(n, F)$ of all $n \times n$ matrices over the field F, introduced in Section 3.5.2, is the endomorphism monoid of a vector space. Also, if X is a non-empty set then the full transformation semigroup T_X is the endomorphism monoid of the (very basic) mathematical object X.

Over the next two chapters we will consider the endomorphism monoids of a number of different mathematical objects. By "mathematical object" what we mean is a general *algebra*, in the sense of universal algebra. The formal definition of an algebra will be given in §4.2. In §4.3 the special class of algebras that we are interested in, namely the so called *independence algebras*, will be defined. In §4.4 we consider generating sets of $\mathrm{End}(\mathcal{A})$ and in §4.5 known results concerning the ideal structure of $\mathrm{End}(\mathcal{A})$ are presented. In §4.6 an outline of the strategy of the proof of the main result is given. In §4.7 and §4.8 endomorphism monoids of finite trivial independence algebras are considered and the main results of the chapter are presented.

4.2 Universal algebra

We will need some basic ideas from the field of universal algebra. In particular, we need the notion of algebra, subalgebra and homomorphism. The main sources for the definitions and results of this section are [27] and [74].

Definition 4.1. Let A be a set and let n be a natural number. A function $f : A^n \to A$ is called an *n-ary operation* defined on A and it is said to have *arity* n.

An n-ary operation f on A can be regarded as an $(n+1)$-ary relation on A (i.e. a subset of A^{n+1}) called the *graph* of f and defined by:

$$\{(a_1, \ldots, a_{n+1}) \in A^{n+1} : f(a_1, \ldots, a_n) = a_{n+1}\}.$$

Example 4.2. Consider the set $\mathbb{N}$ of natural numbers and the binary operation $f : A \times A \to A$ defined by $f(a, b) = a + b$. It has the graph:

$$\{(1,1,2), (1,2,3), (2,1,3), (2,2,4), (2,3,5), (3,2,5), \ldots\}$$

which may be visualised as

$$
\begin{array}{c|cccc}
\vdots & & & & \\
3 & 4 & 5 & 6 & \\
2 & 3 & 4 & 5 & \\
1 & 2 & 3 & 4 & \\
\hline
 & 1 & 2 & 3 & \dots
\end{array}
.
$$

In this way the notion of graph is a generalization of that of Cayley table.

The definition of n-ary operation can be extended to the special case where $n = 0$ to get a, so called, *nullary* operation. In this case we define $A^0 = \{\varnothing\}$ so that $f : \{\varnothing\} \to A$. A nullary operation on A is uniquely determined by $f(\varnothing) \in A$. Conversely, for all $a \in A$ there is exactly one mapping $f_a : \{\varnothing\} \to A$ with $f_a(\varnothing) = a$. Therefore, a nullary operation may be thought of as selecting an element from A.

Definition 4.3. Let A be a non-empty set. Let I be some non-empty index set and let $({f_i}^A)_{i \in I}$ be a function that assigns to each element of I an n_i-ary operation ${f_i}^A$ defined on A. Then the pair $\mathcal{A} = (A\ ;\ ({f_i}^A)_{i \in I})$ is called an *algebra* (indexed by the set I).

The set A is called the *universe* of the algebra $\mathcal{A}$, and $({f_i}^A)_{i \in I}$ is called the *sequence of fundamental operations of* $\mathcal{A}$. We call the number n_i the *arity of* ${f_i}^A$ and the sequence $\tau = (n_i)_{i \in I}$ of all arities is called the *type* of the algebra.

Often f^A will be denoted simply by f. Operations of arities zero, one, two and three are often said to be nullary, unary, binary and ternary operations.

Example 4.4. A monoid $\mathcal{M} = (M\ ;\ \cdot, e)$ is an algebra of type $(2, 0)$ that satisfies the identities

$$
\begin{array}{ll}
\forall x, y, z \in M & x \cdot (y \cdot z) = (x \cdot y) \cdot z \\
\forall x \in M & x \cdot e = e \cdot x = x.
\end{array}
$$

Example 4.5. A ring $\mathcal{R} = (R\ ;\ +, \cdot, -, 0)$ is an algebra of type $(2, 2, 1, 0)$ where $(R\ ;\ +, -, 0)$ is an abelian group, $(R\ ;\ \cdot)$ is a semigroup and

$$
\begin{array}{ll}
\forall x, y, z \in R & x \cdot (y + z) = x \cdot y + x \cdot z \\
\forall x, y, z \in R & (x + y) \cdot z = x \cdot z + y \cdot z.
\end{array}
$$

Example 4.6. Let $\mathcal{R} = (R\ ;\ +, \cdot, -, 0)$ be a ring such that R is infinite. An algebra $\mathcal{M} = (M\ ;\ +, -, 0, R)$ of type $(2, 1, 0, (1)_{r \in R})$, so there is a unary operation for every $r \in R$, is called an $\mathcal{R}$-module if $(M\ ;\ +, -, 0)$ is an abelian group and if for all $r, s \in R$ and $x, y \in M$

$$\begin{aligned} r(x+y) &= r(x) + r(y) \\ (r+s)x &= r(x) + s(x) \\ (r \cdot s)(x) &= r(s(x)). \end{aligned}$$

If R also has an identity element 1 then we must also have $1(x) = x$. In the special case that $\mathcal{R}$ is a field, an $\mathcal{R}$-module is usually called an $\mathcal{R}$ *vector space.*

In this way, modules and vector spaces may be thought of as algebras with infinitely many operations (when $\mathcal{R}$ is infinite).

Example 4.7. Let X be a non-empty set. Then X may be viewed as the algebra $\mathcal{X} = (X\ ;\ )$ which has an empty set of operations.

Our interest will be in semigroups of mappings between algebras. We will need the notion of subalgebra.

Definition 4.8. Let $\mathcal{B} = (B\ ;\ ({f_i}^B)_{i \in I})$ be an algebra of type τ. Then an algebra $\mathcal{A}$ is called a subalgebra of $\mathcal{B}$, written $\mathcal{A} \leq \mathcal{B}$, if

(i) $\mathcal{A} = (A\ ;\ ({f_i}^A)_{i \in I})$ is an algebra of type τ;

(ii) $A \subseteq B$;

(iii) For all $i \in I$ the graph of ${f_i}^A$ is a subset of the graph of ${f_i}^B$.

Note that, in particular, a nullary operation f_i must designate the same element in each subalgebra $\mathcal{A}$ of $\mathcal{B}$.

Proposition 4.9. *The non-empty intersection $\mathcal{A}$ of a non-empty family $\{A_j | j \in J\}$ of subalgebras of an algebra $\mathcal{B}$ is a subalgebra of $\mathcal{B}$.*

Definition 4.10. Let $X \subseteq B$ with $X \neq \varnothing$. Define:

$$\langle X \rangle_{\mathcal{B}} = \bigcap \{\mathcal{A} \mid \mathcal{A} \leq \mathcal{B}\ \&\ X \subseteq A\},$$

the *subalgebra of $\mathcal{B}$ generated by X.*

It is often more convenient to think about generation of subalgebras in the following way. Define

$$E(X) = X \cup \{{f_i}^B(a_1, \ldots, a_{n_i}) \mid i \in I, a_1, \ldots, a_{n_i} \in X\}.$$

Then define inductively

$$E^0(X) = X, \quad E^{k+1}(X) = E(E^k(X))$$

for all $k \in \mathbb{N}$.

Proposition 4.11. *Let $\mathcal{B}$ be an algebra and let X be a non-empty subset of B. Then*

$$\langle X \rangle_{\mathcal{B}} = \bigcup_{k=0}^{\infty} E^k(X).$$

When there is no danger of ambiguity in doing so we write $\langle X \rangle$ rather than $\langle X \rangle_{\mathcal{B}}$ for the subalgebra generated by X.

In a poset we say that z is a *lower bound* of x and y if $z \leq x$ and $z \leq y$. A *greatest lower bound* is a maximal element in the set of lower bounds. Upper bounds and greatest upper bounds may be defined in a similar way. A *lattice* is a poset in which each pair of elements has a unique greatest lower bound and a unique least upper bound. If we take the set of all subalgebras of an algebra $\mathcal{A}$ and order it by inclusion (i.e. $\mathcal{B}_1 \leq \mathcal{B}_2 \Leftrightarrow B_1 \subseteq B_2$) then the resulting poset is a lattice. We call this lattice the *subalgebra lattice* of the algebra $\mathcal{A}$

Finally we need the notion of a homomorphism between algebras.

Definition 4.12. Let $\mathcal{A} = (A\ ;\ ({f_i}^A)_{i \in I})$ and $\mathcal{B} = (B\ ;\ ({f_i}^B)_{i \in I})$ be algebras of the same type. A function $\phi : A \to B$ is called a homomorphism of $\mathcal{A}$ into $\mathcal{B}$ if for all $i \in I$

$$\phi({f_i}^A(a_1, \ldots, a_{n_i})) = {f_i}^B(\phi(a_1), \ldots, \phi(a_{n_i}))$$

for all $a_1, \ldots, a_{n_i} \in A$. In the special case that $n_i = 0$, this equation means that $\phi({f_i}^A(\varnothing)) = {f_i}^B(\varnothing)$. That is, the element designated by the nullary operation ${f_i}^A$ in A must be mapped to the corresponding element ${f_i}^B$ in B.

If the map ϕ is bijective then ϕ is called an *isomorphism.* A homomorphism of an algebra into itself is called an *endomorphism* of $\mathcal{A}$ and a bijective endomorphism is called an *automorphism* of $\mathcal{A}$. The set of all automorphisms of an algebra $\mathcal{A}$ is closed under composition and forms a group that we call the automorphism group of $\mathcal{A}$ and denote $\mathrm{Aut}(\mathcal{A})$. Similarly, the set of endomorphisms of $\mathcal{A}$ is closed under composition and forms a monoid which we denote $\mathrm{End}(\mathcal{A})$. If $\mathcal{A}$ is an algebra we call the subalgebra $\langle \varnothing \rangle$ the *constants* of the algebra. The endomorphisms of an algebra $\mathcal{A}$ must act identically on the set of constants. In other words if $\alpha \in \mathrm{End}(\mathcal{A})$ and $c \in \langle \varnothing \rangle$ then $c\alpha = c$. Also, the image of any endomorphism is a subalgebra of $\mathcal{A}$.

All of the examples of transformation semigroup given in previous chapters may be viewed as endomorphism monoids of various algebras. Indeed, it is easily observed, for example, that each monoid M is isomorphic to its monoid of left translations (Cayley's theorem), which is precisely the monoid of endomorphisms of the algebra $\mathcal{M} = (M \; ; \; (f_m)_{m \in M})$ of type $(1)_{m \in M}$ where f_m is defined by $f_m(s) = ms$, the product of m and s in M.

Example 4.13. Another example is given by the semigroup of order preserving mappings O_n which is the endomorphism monoid of the algebra $\mathcal{Y} = (X_n \; ; \; \cdot)$ of type (2) where the binary operation $\cdot : X_n \times X_n \rightarrow X_n$ is defined by $(a, b) \mapsto \min(a, b)$. Thus for $a \leq b$ and $\phi \in \text{End}(\mathcal{A})$ we have

$$\phi(a) = \phi(\min(a, b)) = \phi(a \cdot b) = \phi(a) \cdot \phi(b) = \min(\phi(a), \phi(b))$$

and so $\phi(a) \leq \phi(b)$ (i.e. the map is order preserving).

4.3 Independence algebras

Looking at Proposition 2.14, which describes the structure of T_n, and Proposition 3.38, which describes the structure of $\text{End}(V)$, there are a striking number of similarities between the two semigroups. In [39] Gould asks the question:

> "What do vector spaces and sets have in common which forces $\text{End}(V)$ and T_n to support a similar pleasing structure?"

She answers this question by defining a class of algebra called an *independence algebra.* In fact the algebras she defines are precisely the v^*-algebras of [76]. An independence algebra is an algebra that satisfies two properties. The first is the *exchange property*:

> [EP] For every subset X of the algebra $\mathcal{A}$ and all elements $x, y \in A$, if $y \in \langle X \cup \{x\} \rangle$ and $y \notin \langle X \rangle$ then $x \in \langle X \cup \{y\} \rangle$.

A subset X of an algebra $\mathcal{A}$ is said to be independent if for all $x \in X$ we have $x \notin \langle X \backslash \{x\} \rangle$ and is called dependent otherwise. There are a number of equivalent ways of stating the exchange property.

Proposition 4.14. *[74, page 50, exercise 6] Let $\mathcal{A}$ be an algebra. Then the following conditions are equivalent.*

(i) The algebra $\mathcal{A}$ satisfies the exchange property.

(ii) For every subset X of $\mathcal{A}$ and for every element u of $\mathcal{A}$ if X is independent and $u \notin \langle X \rangle$ then $X \cup \{u\}$ is independent.

(iii) For every subset X of $\mathcal{A}$ if Y is a maximal independent subset of X then $\langle Y \rangle = \langle X \rangle$.

(iv) For subsets X, Y of $\mathcal{A}$ with $Y \subseteq X$ if Y is independent then there is an independent set Z with $Y \subseteq Z \subseteq X$ and $\langle Z \rangle = \langle X \rangle$. □

A *basis* for $\mathcal{A}$ is a subset of A which generates $\mathcal{A}$ and is independent. Any algebra satisfying [EP] has a basis. Moreover, in such an algebra a subset X is a basis if and only if X is a minimal generating set which is true if and only if X is a maximal independent set. All of the bases of $\mathcal{A}$ have the same cardinality which we call the *dimension* of the algebra. By the third of the equivalent conditions above, it follows that any independent subset of an algebra satisfying [EP] can be extended to a basis.

The second property that an independence algebra must satisfy is the *free basis property*:

> [FB] Any map from a basis of $\mathcal{A}$ into $\mathcal{A}$ can be extended to an endomorphism of $\mathcal{A}$.

Examples of independence algebras include sets (with an empty set of operations), vector spaces and free (right) G-sets (which will be discussed in the next section). In [39] independence algebras are defined and some properties of their endomorphism monoids are given. In the papers [30] and [32] Fountain and Lewin consider the subsemigroup generated by the idempotents of $\mathrm{End}(\mathcal{A})$, firstly for finite dimensional independence algebras, and then for the infinite dimensional case. These results provide a common generalisation of the results of Howie [55] and Erdos [29], in the former case, and of Howie [55] and Reynolds and Sullivan [79], in the latter. Also, in [33] Fountain obtained combinatorial results for the depths of endomorphism monoids $\mathrm{End}(\mathcal{A})$ where $\mathcal{A}$ is a *strong* independence algebra. As far as things stand, however, no generalisation has yet been made of the results on minimal generating sets of ideals of T_n and of $\mathrm{End}(V)$ to the wider class of semigroups $\mathrm{End}(\mathcal{A})$, where $\mathcal{A}$ is an independence algebra. These are the questions that will be addressed in the following two chapters.

Included in [39] is a description of Green's relations in $\mathrm{End}(\mathcal{A})$.

Lemma 4.15. *[39, Corollary 4.6] Let $\mathcal{A}$ be an independence algebra. Then for $\alpha, \beta \in \mathrm{End}(\mathcal{A})$ we have:*

(i) $\alpha\mathcal{L}\beta$ *if and only if* $\operatorname{im}\alpha = \operatorname{im}\beta$*;*

(ii) $\alpha\mathcal{R}\beta$ *if and only if* $\ker(\alpha) = \ker(\beta)$*;*

(iii) $\alpha\mathcal{D}\beta$ *if and only if* $\dim(\operatorname{im}\alpha) = \dim(\operatorname{im}\beta)$*;*

(iv) $\mathcal{J} = \mathcal{D}$. □

In condition (iii), it makes sense to speak of the dimension of $\operatorname{im}\alpha$ because any subalgebra of an independence algebra is again an independence algebra and thus has a well defined dimension. Since $\mathcal{J} = \mathcal{D}$ we will adopt the convention of always referring to the $\mathcal{D}$-classes and not to the $\mathcal{J}$-classes.

For an n-dimensional independence algebra $\mathcal{A}$ the ideals of $\operatorname{End}(\mathcal{A})$ are the sets

$$I_r = \{\alpha \in \operatorname{End}(\mathcal{A}) : \dim(\operatorname{im}\alpha) \leq r\},$$

where $0 \leq r \leq n$, and $I_r = D_0 \cup \cdots \cup D_r$, a union of $\mathcal{D}$-classes, where

$$D_r = \{\alpha \in \operatorname{End}(\mathcal{A}) : \dim(\operatorname{im}\alpha) = r\}$$

and $D_0 \leq D_1 \leq \ldots \leq D_r$. The next result is taken from [30]. It tells us that the minimal (with respect to inclusion) generating sets of I_r are contained wholly within its unique maximal $\mathcal{D}$-class D_r.

Lemma 4.16. *[30, Lemma 2.2] Let* $\alpha \in D_{r-1}$ *where* $1 \leq r \leq n-1$*. Then there are endomorphisms* $\beta, \gamma \in D_r$ *such that* $\alpha = \beta\gamma$*.*

If $\mathcal{A}$ has constants then I_0 is non-empty and is a principal factor in $\operatorname{End}(\mathcal{A})$. If $\mathcal{A}$ does not have constants then I_0 is empty and I_1 is a principal factor. In either case the other principal factors are the remaining Rees quotients I_r/I_{r-1}.

For $r \in \mathbb{N}$ we denote the principal factor I_r/I_{r-1} by P_r. If $\mathcal{A}$ has constants then $P_1 = I_1/I_0$ and $P_0 = I_0$, otherwise $P_1 = I_1$. It was proven in [39] that each P_r is a completely 0-simple semigroup. It follows from Lemma 4.16 that $\operatorname{rank}(I_r) = \operatorname{rank}(P_r)$ and that $\operatorname{idrank}(I_r) = \operatorname{idrank}(P_r)$.

We conclude the section with a method, first described in [14], for building new independence algebras from old. Let $\mathcal{A}$ be an independence algebra and $\mathcal{B}$ be a subalgebra of $\mathcal{A}$. Let $\mathcal{A}[\mathcal{B}]$ denote the algebra $\mathcal{A}$ with the additional nullary operations ν_b, with value b, for all $b \in \mathcal{B}$. If $\mathcal{A}$ is an independence algebra with dimension n and $\mathcal{B}$ is a subalgebra with dimension $m \leq n$ then $\mathcal{A}[\mathcal{B}]$ is an independence algebra with dimension $n - m$. The subalgebras of $\mathcal{A}[\mathcal{B}]$ are the subalgebras of $\mathcal{A}$ that contain $\mathcal{B}$, and the endomorphisms of $\mathcal{A}[\mathcal{B}]$ are the endomorphisms of $\mathcal{A}$ that fix $\mathcal{B}$ elementwise. In other words

$$\mathrm{End}(\mathcal{A}[\mathcal{B}]) = \{\alpha \in \mathrm{End}(\mathcal{A}) : b\alpha = b, \ \forall b \in B\}.$$

4.4 Generating sets for End($\mathcal{A}$)

We begin with the problem of finding generating sets for the semigroup End($\mathcal{A}$). In the special cases of T_n and End(V) the answer is already known. As mentioned in the introduction to Chapter 2 the full transformation semigroup T_n has rank 3. In fact, T_n is generated by S_n, its group of units, together with any element from the $\mathcal{J}$-class J_{n-1}. In a similar way, the general linear semigroup of $n \times n$ matrices over a finite field is generated by three elements (and no fewer). Again, GLS(n, F) is generated by its group of units GL(n, F) together with any matrix with rank $n - 1$. This was originally proven by Waterhouse in [94].

Of course, in general it will not be the case that every finite independence algebra $\mathcal{A}$ satisfies rank(End($\mathcal{A}$)) $= 3$. In fact, for every group G there is an independence algebra $\mathcal{A}$ such that End($\mathcal{A}$) $\cong G$. Namely the algebra $\mathcal{A} = (G\,;\,(f_g)_{g \in G})$ of type $(1)_{g \in G}$, one unary operation for every element of G, where $f_g(h) = gh$. It is easy to see that $\mathcal{A}$ is a one dimensional independence algebra with $\langle \varnothing \rangle = \varnothing$. The bases of this algebra are the singleton subsets of G and End($\mathcal{A}$) $=$ Aut($\mathcal{A}$) $\cong G$. Therefore the question of determining rank(End($\mathcal{A}$)) is at least as hard as the question of determining rank(G) for all finite groups G. The interesting thing about T_n and End(V) is not that the ranks both equal three, but that in order to generate the semigroup, once the group of units has been generated, only one more element needs to be added. In the language of relative ranks, introduced in Section 2.2, if the independence algebra $\mathcal{A}$ is either a finite set or a finite vector space then

$$\mathrm{rank}(\mathrm{End}(\mathcal{A}) : \mathrm{Aut}(\mathcal{A})) = 1.$$

Is the same true of all finite Independence algebras? This is the question we address in this section.

Definition 4.17. Let n and m be natural numbers. Define $P_{n,m}$ to be the subsemigroup of T_{n+m} made up of all maps that fix the last m points, elementwise. In other words:

$$P_{n,m} = \{\alpha \in T_{n+m} : i\alpha = i \text{ for } i = n+1, \ldots, n+m\}.$$

In particular, $P_{n,1} \cong P_n$ (the *partial transformation monoid* on n points). We call $P_{n,m}$ the *generalized partial transformation monoid.*

The monoid $P_{n,m}$ is the endomorphism monoid of the independence algebra $\mathcal{A}[\mathcal{B}]$ where $\mathcal{A} = (X_{n+m}\ ;\ )$ is just a set, and $\mathcal{B} \leq \mathcal{A}$ with $B = \{n+1, \dots, n+m\}$. Thus $\mathcal{A}[\mathcal{B}]$ is the algebra

$$\mathcal{A}[\mathcal{B}] = (\{1, \dots, n+m\}\ ;\ (f_i)_{i \in \{n+1,\dots,n+m\}})$$

of type $(0)_{i \in \{n+1,\dots,n+m\}}$ where the operations are defined by $f_i(\varnothing) = i$. It follows that $\langle \varnothing \rangle = \{n+1, \dots, n+m\}$ and the subalgebras of $\mathcal{A}$ of dimension r are subsets of X_{n+m} of the form $B \cup \{n+1, \dots, n+m\}$ where $B \subseteq X_n$ and $|B| = r$. It follows that the $\mathcal{D}$-classes of $\operatorname{End}(\mathcal{A})$ are given by

$$D_r = \{\alpha \in P_{n,m} : |\operatorname{im} \alpha \cap X_n| = r\}$$

for $0 \leq r \leq n$.

Proposition 4.18. *Let m and n be natural numbers with $n \geq 3$. Then the monoid $P_{n,m}$ satisfies:*

$$\operatorname{rank}(P_{n,m}) = 3 + m.$$

Proof. We begin with some notation. For every $\alpha \in P_{n,m}$ let

$$\sigma(\alpha) = \{1, \dots, n\}\alpha \cap \{n+1, \dots, n+m\}.$$

Observe that for every $\alpha \in D_{n-1}$ either $\sigma(\alpha) = \varnothing$ or $\sigma(\alpha) = \{j\}$ for some $j \in \{n+1, \dots, n+m\}$.

We now show that a generating set of the required size can be found. Identify, in the obvious way, elements of the symmetric group S_n with those of the group of units of $P_{n,m}$. Let $\alpha, \beta \in S_n$ generate S_n and let:

$$\gamma = \left(\begin{array}{ccccc|ccc} 1 & 2 & 3 & \dots & n & n+1 & \dots & n+m \\ 2 & 2 & 3 & \dots & n & n+1 & \dots & n+m \end{array}\right),$$

$$\gamma_j = \left(\begin{array}{ccccc|ccc} 1 & 2 & 3 & \dots & n & n+1 & \dots & n+m \\ j & 2 & 3 & \dots & n & n+1 & \dots & n+m \end{array}\right)$$

for $j = n+1, \dots, n+m$. Let

$$A = \{\alpha, \beta, \gamma\} \cup \{\gamma_i : i = n+1, \dots, n+m\}.$$

Then $|A| = 3 + m$ and we claim that A generates $P_{n,m}$. Let $\delta \in J_{n-1}$. There are two cases to consider depending on the value of $\sigma(\delta)$. If $\sigma(\delta) = \varnothing$ then there exist $\alpha_1, \alpha_2 \in S_n$ such that $\alpha_1\gamma\alpha_2 = \delta$. On the other hand, if $\sigma(\delta) = \{k\}$ then there exist $\beta_1, \beta_2 \in S_n$ such that $\beta_1\gamma_k\beta_2 = \delta$. It follows that $J_{n-1} \subseteq \langle A \rangle$ and, by Lemma 4.16, A generates $P_{n,m}$.

For the converse let A be a generating set for $P_{n,m}$. We claim that $A \cap J_{n-1}$ must contain an element α such that $\sigma(\alpha) = \varnothing$. Indeed, if no such element belonged to A then every $\alpha \in \langle A \rangle \setminus S_n$ would satisfy $\sigma(\alpha) \neq \varnothing$ and as a consequence $\langle A \rangle \neq P_{n,m}$.

Claim. *Let $\gamma_1, \dots, \gamma_r \in P_{n,m}$ and let $\delta = \gamma_1 \dots \gamma_r$. If $i \in \sigma(\delta)$ then $i \in \sigma(\gamma_s)$ for some s.*

Proof. Let t be the smallest number such that $i \in \sigma(\gamma_1\gamma_2 \dots \gamma_t)$. If $t = 1$ then $i \in \sigma(\gamma_1)$. Otherwise there is a $k \in X_n$ such that $k\gamma_1\gamma_2 \dots \gamma_{t-1} \in X_n$ and $(k\gamma_1\gamma_2 \dots \gamma_{t-1})\gamma_t = i$. It follows from the definition of σ that $i \in \sigma(\gamma_t)$. □

It follows from the above claim that for every $j \in \{n+1, \dots, n+m\}$ the set $A \cap J_{n-1}$ must contain some element α such that $\sigma(\alpha) = \{j\}$. Since, including $\varnothing$, there are $m+1$ possibilities for the set $\sigma(\alpha)$ where $\alpha \in J_{n-1}$, we conclude that $|A| \geq 2 + (m+1)$ and so

$$\operatorname{rank}(P_{n,m}) \geq 3 + m.$$

□

Therefore, for every natural number n there exists a finite independence algebra $\mathcal{A}$ that satisfies:

$$\operatorname{rank}(\operatorname{End}(\mathcal{A}) : \operatorname{Aut}(\mathcal{A})) = n.$$

If we consider only independence algebras where $\langle \varnothing \rangle = \varnothing$ then things work out differently.

Theorem 4.19. *Let $\mathcal{A}$ be an n-dimensional independence algebra with $\langle \varnothing \rangle = \varnothing$. If $\alpha \in \operatorname{End}(\mathcal{A})$ is any endomorphism with $\dim(\operatorname{im} \alpha) = n - 1$ then:*

$$\langle \operatorname{Aut}(\mathcal{A}) \cup \{\alpha\} \rangle = \operatorname{End}(\mathcal{A}).$$

In particular

$$\operatorname{rank}(\operatorname{End}(A) : \operatorname{Aut}(A)) = 1.$$

Before proving the result we first need to prove a lemma.

Definition 4.20. Let $\alpha \in \mathrm{End}(\mathcal{A})$. A subset X of $\mathcal{A}$ is a *preimage basis* for α if α is one-one on X and $X\alpha$ is a basis for $\mathrm{im}\,\alpha$.

In [39] it is shown that if X is a preimage basis for $\mathcal{A}$ then X is an independent subset of A. Using this definition we prove the following lemma which describes what the elements of J_{n-1} look like.

Lemma 4.21. *Let $\mathcal{A}$ be an n-dimensional independence algebra with $\langle\varnothing\rangle = \varnothing$ and let $\alpha \in \mathrm{End}(\mathcal{A})$ with $\dim(\mathrm{im}\,\alpha) = n-1$. Then there exists a basis $B = \{b_1, \ldots, b_n\}$ of $\mathcal{A}$ and a basis $Q = \{q_1, \ldots, q_{n-1}\}$ of $\mathrm{im}\,\alpha$ such that:*

$$b_i\alpha = \begin{cases} q_i & \text{if } i \in \{1, \ldots, n-1\} \\ q_{n-1} & \text{if } i = n. \end{cases}$$

Proof. Let $D = \{d_1, \ldots, d_{n-1}\}$ be a basis for $\mathrm{im}\,\alpha$. Let $C = \{c_1, \ldots, c_{n-1}\}$ be a preimage basis of D where $c_i \mapsto d_i$ for all i. Then C is an independent subset of $\mathcal{A}$ and therefore generates an $n-1$ dimensional subalgebra $\langle C\rangle$ of $\mathcal{A}$. Restricting α to the subspace $\langle C\rangle$ gives a homomorphism (in fact an isomorphism) from $\langle C\rangle$ onto $\langle D\rangle$. Now choose and fix an element d belonging to the image of $\mathcal{A}\setminus\langle C\rangle$ under α, say $x \in A \setminus \langle C\rangle$ where $x\alpha = d$. Note that $d \in \mathrm{im}\,\alpha = \langle D\rangle$. Since $d \notin \langle\varnothing\rangle = \varnothing$ we can extend d to a basis $E = \{d, e_2 \ldots, e_{n-1}\}$ of $\mathrm{im}\,\alpha$. If d were allowed to be a constant then this would not be possible. Since α maps $\langle C\rangle$ onto $\langle D\rangle$ we can find a basis $L = \{a_1, \ldots, a_{n-1}\}$ of $\langle C\rangle$ such that $a_1\alpha = d,\ a_2\alpha = e_2,\ \ldots,\ a_{n-1}\alpha = e_{n-1}$. Now $x \notin \langle L\rangle = \langle C\rangle$ and it follows that $L \cup \{x\}$ is a basis for $\mathcal{A}$ and satisfies the conditions given in the statement of the lemma. □

The result above does not hold for independence algebras where $\langle\varnothing\rangle \neq \varnothing$.

Example 4.22. Let $\mathcal{A}$ be the independence algebra with universe $\{1, 2, 3, 4\}$ and a single nullary operation with image 4. Consider the endomorphism

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 4 \end{pmatrix} \in D_2.$$

The unique basis of $\mathcal{A}$ is $\{1, 2, 3\}$ and the unique basis of $\mathrm{im}\,\alpha$ is $\{1, 2\}$. With respect to these bases α does not have the form given in Lemma 4.21 (since $3 \mapsto 4$ under α).

Proof of Theorem 4.19. We show that given $\alpha, \beta \in J_{n-1}$ there exist $\gamma, \delta \in \mathrm{Aut}(\mathcal{A})$ such that

$$\gamma\alpha\delta = \beta.$$

Let $\alpha, \beta \in J_{n-1}$. By Lemma 4.21 there exist bases $B = \{b_1, \ldots, b_n\}$ and $C = \{c_1, \ldots, c_n\}$ of $\mathcal{A}$ such that

$$b_i\alpha = \begin{cases} q_i & \text{if } i \in \{1, \ldots, n-1\} \\ q_{n-1} & \text{if } i = n \end{cases}, \quad c_i\beta = \begin{cases} p_i & \text{if } i \in \{1, \ldots, n-1\} \\ p_{n-1} & \text{if } i = n. \end{cases}$$

Define γ to be the unique endomorphism that satisfies $c_i\gamma = b_i$ for all i. Extend $\{q_1, \ldots, q_{n-1}\}$ by $\{q_n\}$ to a basis of $\mathcal{A}$. Similarly extend $\{p_1, \ldots, p_{n-1}\}$ by $\{p_n\}$ to a basis of $\mathcal{A}$. Let δ be the unique endomorphism satisfying $q_i\delta = p_i$ for all i. Then $\gamma, \delta \in \mathrm{Aut}(\mathcal{A})$ with

$$\begin{aligned} c_i(\gamma\alpha\delta) &= b_i\alpha\delta = \left(\begin{cases} q_i & \text{if } i \in \{1, \ldots, n-1\} \\ q_{n-1} & \text{if } i = n. \end{cases} \right)\delta \\ &= \begin{cases} p_i & \text{if } i \in \{1, \ldots, n-1\} \\ p_{n-1} & \text{if } i = n \end{cases} = c_i\beta \end{aligned}$$

and therefore $\gamma\alpha\delta = \beta$. This fact, along with Lemma 4.16, implies that for every $\alpha \in J_{n-1}$

$$\langle \mathrm{Aut}(\mathcal{A}) \cup \{\alpha\} \rangle = \langle \mathrm{Aut}(\mathcal{A}) \cup J_{n-1} \rangle = \mathrm{End}(\mathcal{A}).$$

□

The converse of Theorem 4.19 does not hold. That is, if

$$\mathrm{rank}(\mathrm{End}(A) : \mathrm{Aut}(A)) = 1$$

it does not follow necessarily that $\langle \varnothing \rangle = \varnothing$. For example, a finite vector space V has the zero vector 0 as a constant but still satisfies $\mathrm{rank}(\mathrm{End}(A) : \mathrm{Aut}(A)) = 1$.

Open Problem 6. Investigate $\mathrm{rank}(\mathrm{End}(A) : \mathrm{Aut}(A))$ in the infinite dimensional case. The answer is known for the full transformation semigroup T_X, in [60] it was shown that if X is infinite then $\mathrm{rank}(T_X : S_X) = 2$.

4.5 Generating sets for ideals of End($\mathcal{A}$)

In [30] it was shown that any proper ideal of End($\mathcal{A}$), where $\mathcal{A}$ is a finite dimensional independence algebra, is idempotent generated. We saw in Theorems 3.35 and 3.39 that, in the cases where $\mathcal{A}$ is a finite set or $\mathcal{A}$ is a finite vector space, not only are the ideals idempotent generated, but they have extremal idempotent generating sets. Does the same hold for all End($\mathcal{A}$) where $\mathcal{A}$ is an arbitrary finite independence algebra? The results of the previous section act as a warning against supposing that the answer must be true just because it is true both for sets and for vector spaces. The main focus of this chapter, and the one that follows it, will be to show that, in most cases, the analogous result does hold for arbitrary finite independence algebras. We will eventually prove:

Theorem 4.23. *Let $\mathcal{A}$ be a finite independence algebra with* $\dim(\mathcal{A}) \geq 3$. *Then every proper two-sided ideal of* $\mathrm{End}(\mathcal{A})$ *has an extremal idempotent generating set.*

In fact we will do more than this. We will go through every class of finite independence algebra given by the classification in [14] and for each class we give a formula for the rank (and idempotent rank) for that class. The condition $\dim(\mathcal{A}) \geq 3$ is necessary. In fact, we will show:

Theorem 4.24. *Let $\mathcal{A}$ be a finite independence algebra with dimension n and let*

$$I_r = \{\alpha \in \mathrm{End}(A) : \dim(\mathrm{im}\,\alpha) \leq r\}$$

where $0 \leq r < n$. Then

$$\mathrm{rank}(I_r) = \max(\rho, \lambda)$$

and

$$\mathrm{idrank}(I_r) = \begin{cases} \mathrm{rank}(I_r) + 1 & \text{if } r = 1 \text{ and } \mathrm{End}(A) \cong P_{2,m} \text{ for some } m \geq 1 \\ \mathrm{rank}(I_r) & \text{otherwise} \end{cases}$$

where ρ and λ are the number of $\mathcal{R}$- and $\mathcal{L}$-classes, respectively, in the unique maximal $\mathcal{J}$-class of I_r.

In light of Lemma 4.16 what we are really interested in are the ranks and idempotent ranks of the completely 0-simple semigroups that appear as principal factors in $\mathrm{End}(A)$. Since the ideals of $\mathrm{End}(\mathcal{A})$ are idempotent generated determining the (ordinary) rank is easy.

Proposition 4.25. *Let $\mathcal{A}$ be a finite independence algebra with dimension n and let*

$$I_r = \{\alpha \in \mathrm{End}(A) : \dim(\mathrm{im}\,\alpha) \leq r\}$$

where $0 \leq r < n$. Then

$$\mathrm{rank}(I_r) = \max(\rho, \lambda)$$

where ρ and λ are the number of $\mathcal{R}$- and $\mathcal{L}$-classes, respectively, in the unique maximal $\mathcal{J}$-class of I_r.

Proof. It follows from Lemma 4.16 that $\text{rank}(I_r) = \text{rank}(P_r)$. Since I_r is idempotent generated, P_r is an idempotent generated completely 0-simple semigroup. By Lemma 3.11 we have $\text{rank}(P_r) = \max(\rho, \lambda)$ and so $\text{rank}(I_r) = \max(\rho, \lambda)$. □

Determining the idempotent rank is less straightforward and is the subject of the rest of this, and the whole of the next, chapter. We will consider only the endomorphism monoids $\text{End}(\mathcal{A})$ where $\mathcal{A}$ has dimension $n \geq 2$. This is reasonable since when $n = 0$ the endomorphism monoid has no proper ideals and when $n = 1$ either there are no proper ideals or there is exactly one which, by [39, Lemma 4.4], must be isomorphic to a left zero semigroup with m^n elements, where m is the number of constants in the algebra. In this case Theorem 4.24 follows trivially from the observation that every band has idempotent rank equal to rank.

4.6 General strategy

We will make use of the classification of finite independence algebras given in [14]. In fact, independence algebras had previously been classified many years ago by those working with v^*-algebras (see [92] for details). Cameron and Szabó's classification [14] is given up to equivalence, where two independence algebras are called equivalent if their subalgebra lattices are isomorphic, and their endomorphism monoids are isomorphic. More precisely, two independence algebras $\mathcal{A}_1$ and $\mathcal{A}_2$ are equivalent if there exists a bijection $\theta : \mathcal{A}_1 \to \mathcal{A}_2$ such that:

(i) both θ and θ^{-1} map subalgebras to subalgebras;

(ii) if $f_i : \mathcal{A}_i \to \mathcal{A}_i$ (for $i = 1, 2$) are maps satisfying $f_1\theta = \theta f_2$ then f_1 is an endomorphism of $\mathcal{A}_1$ if and only if f_2 is an endomorphism of $\mathcal{A}_2$.

The main theorem of [14] states that every finite independence algebra is equivalent to one of the classes that can be found in Sections 2 and 3 of the paper. Over the next four sections we will go through the classes of algebra in those two sections describing the endomorphism monoids, and determining the rank and idempotent rank, in each case.

There are effectively four kinds of independence algebra to consider, depending on whether or not the algebra is *trivial* (a class of algebra that we define at the beginning of the next section) and whether or not it has constants. In each case we follow the four steps described below.

(I) Find a concrete way of representing the endomorphism monoid, usually as a semigroup of pairs.

(II) Describe Green's relations in the monoid, identify the group $\mathcal{H}$-classes and determine the dimensions of the $\mathcal{D}$-classes.

(III) For each $\mathcal{D}$-class, D_r, use the position of the group $\mathcal{H}$-classes to construct a rectangular 0-band T that satisfies $\text{idrank}(T) = \text{idrank}(I_r)$.

(IV) Compute the idempotent rank of T using the results of Chapter 3 along with Theorem 3.35 and Theorem 3.39.

This approach is made clearer by thinking in terms of egg-box pictures. For each class of independence algebra we get enough information to draw an egg-box picture of $\text{End}(\mathcal{A})$. We compare this picture with the egg-box pictures of T_n and $\text{End}(V)$. In each case the picture of $\text{End}(\mathcal{A})$ contains copies of, in the sense described in Section 3.2, either the egg-box picture of T_n or that of $\text{End}(V)$. In the majority of cases the existence of these substructures is all that we need in order to find the idempotent rank.

4.7 Trivial independence algebras without constants

Let G be a group. A *right G-set* is a triple (X, G, ψ) where $\psi : X \times G \to X$ is a map, written $(x, g)\psi = x \cdot g \in X$, that satisfies

(i) $(x \cdot g) \cdot h = x \cdot (gh)$

(ii) $x \cdot 1 = x$

for all $x \in X$ and $g, h \in G$. We say that the group G acts on the set X. It is customary to say that X is a right G-set, without giving a name to the homomorphism. Let G be a group and I be some index set (most of the time we are concerned with I finite and we will usually suppose $I = \{1, \ldots, n\}$). The free right G-set over I consists of pairs (i, g) with the action $(i, g) \cdot h = (i, gh)$.

An independence algebra is called *trivial* if its subalgebra lattice is isomorphic to the power set $\mathcal{P}(X)$ (ordered by inclusion) for some set X. In [14] the authors describe all the trivial independence algebras, up to equivalence, in terms of group actions. Let X be a set, G a group and C a right G-set. The set $(X \times G) \cup C$ forms an independence algebra

$$\mathcal{A} = ((X \times G) \cup C \; ; \; (\nu_c)_{c \in C}, \; (\rho_g)_{g \in G})$$

of type $((0)_{c\in C},(1)_{g\in G})$ where $(\varnothing)\nu_c = c$ and

$$\begin{aligned} (x,h)\rho_g &= (x,hg), \quad x\in X, g,h\in G \\ (c)\rho_g &= c\cdot g, \qquad\quad c\in C. \end{aligned}$$

This algebra has dimension $|X|$, the subalgebras are the subsets of the form $(Y\times G)\cup C$ for Y a subset of X, and $\langle\varnothing\rangle = C$.

We begin by considering trivial independence algebras without constants. As a result of the above classification, trivial independence algebras without constants are just free (right) G-sets. The endomorphism monoids of free and projective S-acts where S is a semigroup (of which free G-sets are an example) were considered in [12]. In that paper the author makes use of a wreath product construction to describe the endomorphism monoid. We will use the same construction here.

We start by describing the construction. It depends on which way round we compose our functions. There are many different functions involved and to try and keep things as simple as possible we view all our functions as mapping on the right and we compose them from left to right always.

Let $\mathcal{A} = (X\times G\ ;\ (\rho_g)_{g\in G})$ be a trivial independence algebra with $\langle\varnothing\rangle = \varnothing$. Any transversal of the copies of G is a basis of this algebra in particular $B = \{(x,1) : x\in X\}$ is a basis that we call the *natural basis* of this algebra. The free basis property tells us that the endomorphisms of $\mathrm{End}(\mathcal{A})$ are uniquely determined by their action on this basis. Associated with every $\pi\in\mathrm{End}(\mathcal{A})$ there are two functions. Firstly we have $\alpha_\pi\in T(X)$ (the full transformation semigroup on X) and secondly we have $f_\pi : X\to G$ defined by:

$$(x,1)\pi = (x\alpha_\pi, xf_\pi).$$

Now consider what happens to the associated functions when we compose two endomorphisms from $\mathrm{End}(\mathcal{A})$. Let $\pi_1,\pi_2\in\mathrm{End}(\mathcal{A})$ and consider the product $\pi_1\pi_2$. We have:

$$\begin{aligned} (x,1)\pi_1\pi_2 &= ((x,1)\pi_1)\pi_2 = (x\alpha_{\pi_1}, xf_{\pi_1})\pi_2 \\ &= ((x\alpha_{\pi_1},1)\cdot(xf_{\pi_1}))\pi_2 = ((x\alpha_{\pi_1},1)\pi_2)\cdot(xf_{\pi_1}) \\ &= ((x\alpha_{\pi_1})\alpha_{\pi_2},(x\alpha_{\pi_1})f_{\pi_2})\cdot(xf_{\pi_1}) = (x\alpha_{\pi_1}\alpha_{\pi_2},(x\alpha_{\pi_1})f_{\pi_2}(xf_{\pi_1})). \end{aligned}$$

This motivates the following definition.

Definition 4.26. Define the semigroup $G\wr T_n$ to be made up of pairs (α, f) where

$\alpha \in T_n$ and $f \in \mathrm{Map}(X_n, G)$ with products defined by:

$$(\alpha, \psi)(\beta, \phi) = (\alpha\beta, \phi^\alpha \psi)$$

where $i\phi^\alpha\psi = ((i\alpha)\phi)\ (i\psi)$ (a product of two elements in G).

To verify that $G \wr T_n$ and $\mathrm{End}(\mathcal{A})$ are isomorphic, define $\xi : \mathrm{End}(A) \to G \wr T_n$ by $\pi \mapsto (\alpha_\pi, f_\pi)$. It is clear that ξ is a bijection. We want to show it is a homomorphism. Since:

$$(i, 1)\pi_1\pi_2 = (i\alpha_{\pi_1}\alpha_{\pi_2}, (i\alpha_{\pi_1})f_{\pi_2}(if_{\pi_1}))$$

it follows that

$$(\pi_1\pi_2)\xi = (\alpha_{\pi_1}\alpha_{\pi_2}, f_{\pi_2}^{\alpha_{\pi_1}} f_{\pi_1}).$$

Therefore

$$(\pi_1\xi)(\pi_2\xi) \;\; = \;\; (\alpha_{\pi_1}, f_{\pi_1})(\alpha_{\pi_2}, f_{\pi_2}) = (\alpha_{\pi_1}\alpha_{\pi_2}, f_{\pi_2}^{\alpha_{\pi_1}} f_{\pi_1}) = (\pi_1\pi_2)\xi$$

and ξ is indeed a homomorphism.

From now on we will view $\mathrm{End}(A)$, where $\mathcal{A}$ is a finite trivial independence algebra without constants, always in terms of the isomorphic semigroups $G \wr T_n$. We start by describing Green's relations in $G \wr T_n$.

Lemma 4.27. *Green's relations in $G \wr T_n$ are given by:*

(i) $(\alpha, \psi)\mathcal{R}(\beta, \phi)$ *if and only if* $\alpha\mathcal{R}\beta$ *in* T_n *and* $(i\phi)(i\psi)^{-1} = (j\phi)(j\psi)^{-1}$ *for all* $(i, j) \in \ker(\alpha)$.

(ii) $(\alpha, \psi)\mathcal{L}(\beta, \phi)$ *if and only if* $\alpha\mathcal{L}\beta$ *in* T_n.

(iii) $(\alpha, \psi)\mathcal{D}(\beta, \phi)$ *if and only if* $|\operatorname{im}\alpha| = |\operatorname{im}\beta|$.

Proof. (i)($\Rightarrow$) If follows from the definition of multiplication in $G \wr T_n$ that $\alpha\mathcal{R}\beta$. For the second part suppose that

$$(\alpha, \psi)(\gamma, \mu) = (\beta, \phi), \quad (\beta, \phi)(\delta, \xi) = (\alpha, \psi).$$

Then we have $\alpha\gamma = \beta$, $\beta\delta = \alpha$, $\mu^\alpha\psi = \phi$ and $\xi^\beta\phi = \psi$ which means that for all $i \in X_n$

$$((i\alpha)\mu)(i\psi) = i(\mu^\alpha\psi) = (i\phi), \quad ((i\beta)\xi)(i\phi) = (\xi^\beta\phi) = (i\psi).$$

Hence for $(i,j) \in \ker\alpha$

$$(i\phi)(i\psi)^{-1} = (i\alpha)\mu = (j\alpha)\mu = (j\phi)(j\psi)^{-1}$$

as required.

($\Leftarrow$) Suppose that the two conditions in (i) hold. Say $\gamma, \delta \in T_n$ with $\alpha\gamma = \beta$ and $\beta\delta = \alpha$. Now define $\mu : X_n \to G$ by:

$$i\mu = \begin{cases} (j\phi)(j\psi)^{-1} & \text{if } i = j\alpha \text{ for some } j \in X_n \\ 1_G & \text{otherwise.} \end{cases}$$

This map is well defined since by assumption if $(i,j) \in \ker\alpha$ then $(i\phi)(i\psi)^{-1} = (j\phi)(j\psi)^{-1}$. Now we have

$$i(\mu^\alpha\psi) = ((i\alpha)\mu)(i\psi) = (i\phi)(i\psi)^{-1}(i\psi) = i\phi$$

implying that $(\alpha,\psi)(\gamma,\mu) = (\beta,\phi)$. Similarly we can find $\xi \in \mathrm{Map}(X_n, G)$ such that $(\beta,\phi)(\delta,\xi) = (\alpha,\psi)$. We conclude that $(\alpha,\psi)\mathcal{R}(\beta,\phi)$ in $G \wr T_n$.

(ii) ($\Rightarrow$) Follows from the definition of multiplication in $G \wr T_n$.

($\Leftarrow$) Suppose that $\delta\alpha = \beta$ and that $\gamma\beta = \alpha$. Then define $\mu, \nu : X_n \to G$ by:

$$i\mu = ((i\delta)\psi)^{-1}(i\phi), \quad i\nu = ((i\gamma)\phi)^{-1}(i\psi).$$

It follows that

$$i(\psi^\delta\mu) = ((i\delta)\psi)(i\mu) = ((i\delta)\psi)((i\delta)\psi)^{-1}(i\phi) = i\phi$$

and

$$i(\phi^\gamma\nu) = ((i\gamma)\phi)(i\nu) = ((i\gamma)\phi)((i\gamma)\phi)^{-1}(i\psi) = i\psi.$$

We conclude that

$$(\delta,\mu)(\alpha,\psi) = (\beta,\phi), \quad (\gamma,\nu)(\beta,\phi) = (\alpha,\psi)$$

and $(\alpha,\psi)\mathcal{L}(\beta,\phi)$ in $G \wr T_n$ as required.

(iii) Follows from Lemma 4.15. □

It follows that for this monoid

$$D_r = \{(\alpha,\psi) \in G \wr T_n : |\operatorname{im}\alpha| = r\}$$

for $1 \leq r \leq n$. We are interested in the distribution of the idempotents in D_r. The following lemma tells us how to identify the group $\mathcal{H}$-classes (and thus find the idempotents).

Lemma 4.28. *Let $(\beta, \psi) \in G \wr T_n$. Then $H_{(\beta,\psi)}$ is a group $\mathcal{H}$-class in $G \wr T_n$ if and only if H_β is a group $\mathcal{H}$-class in T_n.*

Proof. If $H_{(\beta,\psi)}$ is a group then $(\beta, \psi)^2 \mathcal{H} (\beta, \psi)$ in $G \wr T_n$ which, by Lemma 4.27, means that $\beta^2 \mathcal{H} \beta$ in T_n and so H_β is a group $\mathcal{H}$-class.

For the converse suppose that H_β is a group $\mathcal{H}$-class in T_n. This is the same thing as saying that $\beta^2 \mathcal{H} \beta$. We have $(\beta, \psi)^2 = (\beta^2, \psi^\beta \psi)$ where $i\psi^\beta\psi = ((i\beta)\psi)(i\psi)$ in G. If $(i, j) \in \ker(\beta)$ then

$$\begin{aligned}(i\psi^\beta\psi)(i\psi)^{-1} &= ((i\beta)\psi)(i\psi)(i\psi)^{-1} = (i\beta)\psi = (j\beta)\psi \\ &= ((j\beta)\psi)(j\psi)(j\psi)^{-1} = (j\psi^\beta\psi)(j\psi)^{-1}.\end{aligned}$$

It follows from Lemma 4.27 that $(\beta, \psi)^2 \mathcal{H} (\beta, \psi)$ and therefore $H_{(\beta,\psi)}$ is a group $\mathcal{H}$-class in $G \wr T_n$. □

We now find an embedding of T_n in $G \wr T_n$ that dictates the structure of $G \wr T_n$. Define $\iota \in \mathrm{Map}(X_n, G)$ to be the map that sends every element of X_n to the identity of G (i.e. $x\iota = 1_G$ for all $x \in X_n$). Let

$$N = \{(\alpha, \iota) : \alpha \in T_n\}$$

noting that N is a subsemigroup of $G \wr T_n$ and is isomorphic to T_n. Also define

$$\mathcal{G} = \{(1, \psi) : \psi \in \mathrm{Map}(X_n, G)\}$$

where 1 denotes the identity of the monoid T_n. The set $\mathcal{G}$ is a subgroup of the group of units of $G \wr T_n$ and is isomorphic to the direct product of n copies of G. The group $\mathcal{G}$ acts on the $\mathcal{R}$-classes of D_r in the following natural way. Let R be an $\mathcal{R}$-class in D_r and let $(1, \psi) \in \mathcal{G}$. Then define

$$(1, \psi) \cdot R = (1, \psi)R = \{(1, \psi)(\beta, \phi) : (\beta, \phi) \in R\}.$$

This is a well defined action of $\mathcal{G}$ on the $\mathcal{R}$-classes in D_r since given any $(1, \psi) \in \mathcal{G}$ and any $\mathcal{R}$-class R in D_r, by Lemma 4.27, the set $(1, \psi)R$ is an $\mathcal{R}$-class in D_r. Given $(\alpha, \psi) \in G \wr T_n$ we use $\mathcal{O}_{(\alpha,\psi)}$ to denote the orbit of the $\mathcal{R}$-class $R_{(\alpha,\psi)}$.

We now prove that the $\mathcal{R}$-classes of $N \cap D_r$ form a transversal of the orbits of the $\mathcal{R}$-classes of D_r under the action of $\mathcal{G}$, that all the orbits have the same size, and that any two $\mathcal{R}$-classes in the same orbit look the same, in terms of the position of idempotents (i.e. that the map $f : R \to (1,\psi)R$ defined by $f(r) = (1,\psi)r$ sends group $\mathcal{H}$-classes to group $\mathcal{H}$-classes).

Lemma 4.29. *The $\mathcal{R}$-classes of $N \cap D_r$ form a transversal of the orbits of the $\mathcal{R}$-classes of D_r.*

Proof. First we show that every orbit contains at least one $\mathcal{R}$-class from $N \cap D_r$. Let $(\alpha,\psi) \in D_r$ and define $\psi^{-1} \in \mathrm{Map}(X_n, G)$ by $i\psi^{-1} = (i\psi)^{-1}$ for all $i \in X_n$. Then $(1,\psi^{-1})(\alpha,\psi) = (\alpha, \psi^1\psi^{-1})$ which belongs to $N \cap D_r$ since for all $i \in X_n$ we have $i\psi^1\psi^{-1} = (i\psi)(i\psi^{-1}) = 1$. It follows that $\mathcal{O}_{(\alpha,\psi)} = \mathcal{O}_{(\alpha,\psi^1\psi^{-1})}$ where $(\alpha,\psi^1\psi^{-1}) \in N \cap D_r$.

To see that only one $\mathcal{R}$-class of $N \cap D_r$ belongs to each orbit, first let $(\alpha,\psi), (\beta,\phi) \in N \cap D_r$ where $R_{(\alpha,\psi)}$ and $R_{(\beta,\phi)}$ belong to the same orbit. Then there exists some $(\beta',\phi') \in R_{(\beta,\phi)}$ and $(1,\varphi) \in \mathcal{G}$ such that

$$(1,\varphi)(\alpha,\psi) = (\beta',\phi').$$

But by definition we have:

$$(1,\varphi)(\alpha,\psi) = (\alpha,\psi^1\varphi)$$

and so $(\beta',\phi') = (\alpha,\psi^1\varphi)$. We conclude that $\alpha = \beta'$ and so $\alpha = \beta'\mathcal{R}\beta$ in T_n. Moreover, for all $i \in X_n$ we have $i\psi = 1 = i\phi$ so in particular for $(i,j) \in \ker\alpha$ we have $i\psi = 1 = j\phi$. It now follows from Lemma 4.27 that $(\alpha,\psi)\mathcal{R}(\beta,\phi)$ and so $R_{(\alpha,\psi)} = R_{(\beta,\phi)}$. □

Lemma 4.30. *Let $(\alpha,\psi) \in G \wr T_n$ and let $(1,\theta) \in \mathcal{G}$. Then $H_{(\alpha,\psi)}$ is a group if and only if $H_{(1,\theta)(\alpha,\psi)}$ is a group.*

Proof. From Lemma 4.28 it follows that $H_{(\alpha,\psi)}$ is a group if and only if H_α is a group in T_n which is true if and only if $H_{(1,\theta)(\alpha,\psi)} = H_{(\alpha,\psi^1\theta)}$ is a group in $G \wr T_n$. □

Lemma 4.31. *For every $(\alpha,\psi) \in D_r$ we have $|\mathcal{O}_{(\alpha,\psi)}| = |G|^{n-r}$.*

Proof. Let $(\alpha,\psi) \in N \cap D_r$. From the definition of the action we have:

$$\mathcal{O}_{(\alpha,\psi)} = \{R_{(\alpha,\phi)} : \phi \in \mathrm{Map}(X_n, G)\}.$$

Fix $(\alpha,\psi) \in D_r$ and define $Y = \{(\alpha,\phi) : (\alpha,\phi)\mathcal{R}(\alpha,\psi)\}$ where α is fixed. We claim that $|Y| = |G|^r$. Indeed, by Lemma 4.27, $(\alpha,\phi)\mathcal{R}(\alpha,\psi)$ if and only if for

$(i,j) \in \ker\alpha$ we have

$$(i\phi) = (j\phi)(j\psi)^{-1}(i\psi).$$

We now count the number of choices for ϕ that satisfy this equation. For each kernel class of α, once the image under ϕ of a single element of this class is determined, the equation above determines the images under ϕ of every other element in the same kernel class. Therefore, if we fix a transversal of the kernel classes of α, once the images of this set under ϕ are defined, the map ϕ will be completely determined. Since the elements in this set can be mapped anywhere in G, there are $|G|^r$ possibilities in total. We conclude that $|\mathcal{O}_{(\alpha,\psi)}| = |G|^n/|G|^r = |G|^{n-r}$. □

The correspondence that the three preceding lemmas establish between the structure of T_n and that of $G \wr T_n$ is demonstrated by the following example.

Example 4.32. Let $G = S_2$, the symmetric group of degree 2. The semigroup $S = G \wr T_3$ may be embedded in T_6 where it is generated by the following transformations:

$$\alpha_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 1 & 2 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 1 & 2 & 5 & 6 \end{pmatrix},$$

$$\alpha_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 4 & 5 & 6 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 1 & 2 & 5 & 6 \end{pmatrix}.$$

The elements α_i generate the group of units $S_2 \wr T_3$ and β is a singular idempotent in the second top $\mathcal{D}$-class of the semigroup. The semigroup has size 216. The group of units is isomorphic to $S_2 \wr S_3$ and has size 48. The group $\mathcal{H}$-classes in the middle $\mathcal{D}$-class are isomorphic to $S_2 \wr S_2$ while the group $\mathcal{H}$-classes in the bottom $\mathcal{D}$-class are isomorphic to $S_2 \wr S_1$.

The egg-box diagram of the semigroup is given in Figure 4.1 and shows how each of the $\mathcal{D}$-classes looks like a "vertical tiling" of the $\mathcal{D}$-classes of the full transformation semigroup. In the language of Section 3.2 this vertical tiling is a direct product (with amalgamated zero) of a principal factor of T_n with a left zero semigroup.

Finally we note that the group $\mathcal{H}$-classes of $G \wr T_n$ in D_r are, as one would expect, isomorphic to $G \wr S_r$. Combining these observations together we conclude:

Proposition 4.33. *Let G be a finite group and let $n \in \mathbb{N}$. Let*

$$K'(n,r) = \{(\alpha,\psi) \in G \wr T_n : |\operatorname{im}\alpha| \leq r\}, \quad K(n,r) = \{\alpha \in T_n : |\operatorname{im}\alpha| \leq r\}.$$

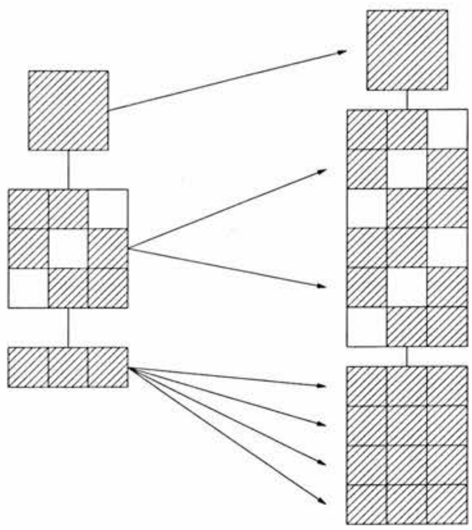


Figure 4.1: Comparing the egg-box pictures of T_3 and $S_2 \wr T_3$. As usual, the shaded boxes correspond to group $\mathcal{H}$-classes.

Also let $V_{n,r} = K(n,r)/K(n,r-1)$, $Q_{n,r} = K'(n,r)/K'(n,r-1)$ *and*

$$D_r = \{(\alpha, \psi) \in G \wr T_n : |\operatorname{im}\alpha| = r\}.$$

Then we have:

(i) The number of $\mathcal{L}$-classes in D_r is $\binom{n}{r}$.

(ii) The number of $\mathcal{R}$-classes in D_r is $|G|^{n-r}S(n,r)$.

(iii) Let T_1 and T_2 be the rectangular 0-band homomorphic images of $Q_{n,r}$ and $V_{n,r}$ respectively. Then $T_1 \cong {L^0}_{|G|^{n-r}} \times_0 T_2$ where ${L^0}_k$ denotes the k-element left zero semigroup with a zero adjoined.

Proof. (i) Follows from Lemma 4.27(ii). (iii) Follows from Lemmas 4.29, 4.30 and 4.31. (ii) Follows from (iii) and Proposition 2.14 which gives the dimensions of the $\mathcal{D}$-classes of T_n.

□

Using the proposition above we now compute the idempotent rank of the ideals of $G \wr T_n$.

Theorem 4.34. *Let G be a finite group and let $n \in \mathbb{N}$. Let*

$$K'(n,r) = \{(\alpha, \psi) \in G \wr T_n : |\operatorname{im}\alpha| \leq r\}$$

where $1 \le r \le n$. Then for $2 \le r < n$ we have

$$\mathrm{idrank}(K'(n,r)) = \mathrm{rank}(K'(n,r)) = |G|^{n-r}S(n,r).$$

Also, when $r = 1$ we have

$$\mathrm{idrank}(K'(n,1)) = \mathrm{rank}(K'(n,1)) = \begin{cases} n & \text{if } |G| = 1 \\ |G|^{n-1} & \text{otherwise.} \end{cases}$$

Proof. For the first part, using Proposition 4.33(iii), and the fact that $r \ge 2$, it follows that the rectangular 0-band T_1 is the direct product with amalgamated zero of the two tall rectangular 0-bands T_2 and ${L^0}_{|G|^{n-r}}$. The rectangular 0-band T_2 has an extremal idempotent generating set by Theorem 3.35. The semigroup ${L^0}_{|G|^{n-r}}$ is a band and so has an extremal idempotent generating set. By Lemma 3.6 it follows that T_1 has an extremal idempotent generating set and, as a consequence of Lemma 4.16, so does $K'(n,r)$.

For the second part we note that, by Proposition 4.33(iii), the $\mathcal{D}$-class in question is completely simple with dimensions $|G|^{n-1} \times n$. The result then follows since $\max(|G|^{n-1}, n)$ is n if G is trivial and $|G|^{n-1}$ otherwise. □

4.8 Trivial independence algebras with constants

These algebras are given by the definition at the beginning of §4.7 with the condition that $C \neq \varnothing$. Let

$$\mathcal{A} = ((X_n \times G) \cup C \ ; \ (\lambda_g)_{g \in G}, \ (\nu_c)_{c \in C})$$

where $C = \{c_1, \ldots, c_m\}$. The bases of this algebra are the transversals of the copies of G. In particular $B = \{(i, 1_G) : i \in X_n\}$ is a basis for the independence algebra $\mathcal{A}$ and every $\alpha \in \mathrm{End}(\mathcal{A})$ is determined by its action on this basis.

We begin by observing that the semigroup $G \wr T_n$ embeds in $\mathrm{End}(\mathcal{A})$ in a particularly nice way. In fact, the way that $G \wr T_n$ embeds in $\mathrm{End}(\mathcal{A})$ is analogous to the way that T_n embeds into P_n.

Lemma 4.35. *Let $T = \{\alpha \in \mathrm{End}(\mathcal{A}) : B\alpha \cap C = \varnothing\}$. Then $T \le \mathrm{End}(\mathcal{A})$ and $T \cong G \wr T_n$.*

Proof. First we show that T is a subsemigroup. Let $\alpha, \beta \in T$. Since $\beta \in T$ it follows that $B\beta \cap C = \varnothing$. It follows that for all $(i,g) \in X_n \times G$ we have $(i,g)\beta \notin C$ (since if $(i,g)\beta \in C$ then $(i, 1_G)\beta = (i, gg^{-1})\beta = ((i,g)\beta) \cdot g^{-1} \in C$). Therefore, if $\{c\} = b\alpha\beta \cap C$ for some $b \in B$ then $b\alpha \in C$ which contradicts the fact that $\alpha \in T$.

To see that $T \cong G \wr T_n$ we just have to show that $T \cong \mathrm{End}(X_n \times G)$, since we have already seen that $\mathrm{End}(X_n \times G) \cong G \wr T_n$. Define $\Psi : T \to \mathrm{End}(X_n \times G)$, noting that B is both a basis for $\mathcal{A}$ and for $X_n \times G$, by

$$b(\alpha\Psi) = b\alpha$$

for $b \in B$. Then the map Ψ is an isomorphism. □

Now consider the $\mathcal{D}$-classes of $\mathrm{End}(\mathcal{A})$. As already mentioned, the subalgebras of $\mathcal{A}$ are the sets $(Y \times G) \cup C$ where $Y \subseteq X$. Therefore, by Theorem 4.15, the $\mathcal{D}$-classes are given by

$$D_r = \{\alpha \in \mathrm{End}(\mathcal{A}) : |\{i \in X_n : (i, g) = b\alpha \text{ for some } b \in B\}| = r\}$$

for $0 \leq r \leq n$. Also by Theorem 4.15 the number of $\mathcal{L}$-classes in D_r, for $1 \leq r \leq n$, is equal to $\binom{n}{r}$, which is the number of r-dimensional subalgebras of $\mathcal{A}$.

Lemma 4.36. *Let $1 \leq r \leq n$. Then $T \cap D_r$ is a union of the $\mathcal{R}$-classes of D_r in* $\mathrm{End}(\mathcal{A})$.

Proof. Let $\tau \in T \cap D_r$ and $\sigma \in \mathrm{End}(\mathcal{A})$ be $\mathcal{R}$-related to τ in $\mathrm{End}(\mathcal{A})$. By Theorem 4.15 it follows that $\ker(\sigma) = \ker(\tau)$. Since $\tau \in T$ it follows that $B\tau \cap C = \varnothing$ and so $\ker(\tau)$ has the set $\{c\}$ among its kernel classes for each $c \in C$. If $\sigma \notin T$ then for some $b \in B$ we would have $b\sigma = c \in C$. But then b and c would belong to the same kernel class of σ and thus $\ker(\sigma) \neq \ker(\tau)$ which is a contradiction. □

It follows from these observations that the $\mathcal{D}$-classes of $G \wr T_n$ "sit inside" the $\mathcal{D}$-classes of $\mathrm{End}(\mathcal{A})$ in the following sense.

Lemma 4.37. *Let G be a finite group and let $m, n \in \mathbb{N}$. Let*

$$K(n, r) = \{(\alpha, \psi) \in G \wr T_n : |\operatorname{im} \alpha| \leq r\}$$

and

$$K'(n, r) = \{\alpha \in \mathrm{End}(\mathcal{A}) : |\{i \in X_n : b\alpha = (i, g) \textit{ for some } b \in B\}| \leq r\}$$

for $1 \leq r \leq n$. Also let

$$V_{n,r} = K(n, r)/K(n, r-1), \quad Q_{n,r} = K'(n, r)/K'(n, r-1).$$

Then $V_{n,r}$ is an $\mathcal{L}$-class filling subsemigroup of $Q_{n,r}$. □

Furthermore, since the group $\mathcal{H}$-classes in any $\mathcal{D}$-class of $\mathrm{End}(\mathcal{A})$ are all isomorphic, and since we know what the group $\mathcal{H}$-classes of $G \wr T_n$ are, we have the following result.

Lemma 4.38. *Let $1 \le r \le n$. Then in $D_r \subseteq \mathrm{End}(\mathcal{A})$ the group $\mathcal{H}$-classes are isomorphic to $G \wr S_r$.*

We now divide our analysis of $\mathrm{End}(\mathcal{A})$ into the case where G is trivial and the case where G is non-trivial.

Trivial independence algebras where $\langle \varnothing \rangle \neq \varnothing$ and G is trivial

When G is trivial, and $|C| = m$, the semigroup $\mathrm{End}(\mathcal{A})$ is isomorphic to the generalized partial transformation semigroup $P_{n,m}$ and the ideals are the sets

$$I(r, n, m) = \{\alpha \in P_{n,m} : |\operatorname{im} \alpha \cap X_n| \le r\}$$

where $1 \le r < n$. It is not true that the ideals of $P_{n,m}$ always have extremal idempotent generating sets as we now show.

Lemma 4.39. *For all $m \in \mathbb{N}$ the semigroup $I(1, 2, m)$ satisfies*

$$\mathrm{idrank}(I(1,2,m)) = \mathrm{rank}(I(1,2,m)) + 1.$$

Proof. The unique maximal $\mathcal{J}$-class of the semigroup $I(1, 2, m)$ is

$$D = \{\alpha \in P_{2,m} : (\{1\alpha\} \cup \{2\alpha\}) \cap \{1, 2\} \neq \varnothing\}.$$

This $\mathcal{J}$-class has only two $\mathcal{L}$-classes which correspond to the elements with images $\{1, 3, \ldots, m+2\}$ and $\{2, 3, \ldots, m+2\}$, respectively. The number of $\mathcal{R}$-classes in D is $2m + 1$ and every $\mathcal{R}$-class, except one, contains only one idempotent. The single exception is the $\mathcal{R}$-class that contains the pair of idempotents ϵ_1, ϵ_2 given by:

$$i\epsilon_1 = \begin{cases} 1 \text{ if } i = 2 \\ i \text{ otherwise,} \end{cases} \qquad i\epsilon_2 = \begin{cases} 2 \text{ if } i = 1 \\ i \text{ otherwise.} \end{cases}$$

Let E be the set of idempotents in D. We claim that any idempotent generating set of $I(1, 2, m)$ must contain all the elements of E. Indeed, we need all of the idempotents $E \backslash \{\epsilon_1, \epsilon_2\}$ since otherwise one of the $\mathcal{R}$-classes of D would be missed

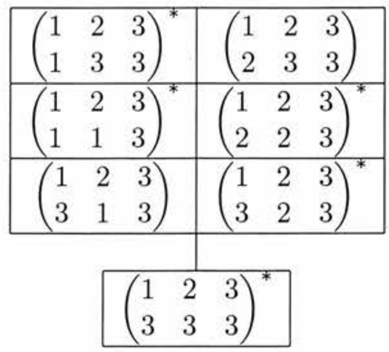


Figure 4.2: The structure of the unique maximal $\mathcal{D}$-class of the semigroup $I(1,2,1)$.

out. Also, since $\epsilon_1 \notin \langle E \cup \{\epsilon_2\}\rangle$ and $\epsilon_2 \notin \langle E \cup \{\epsilon_1\}\rangle$ we conclude that both ϵ_1 and ϵ_2 are required in any idempotent generating set. Thus

$$\text{idrank}(I(1,2,m)) = |E| = 2m+2 = \text{rank}(I(1,2,m)) + 1$$

as required. $\square$

Example 4.40. Consider $I(1,2,1)$ which is one of the two sided ideals of $P_{2,1}$. The elements are

$$I(1,2,1) = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 3 \end{pmatrix}, \right.$$
$$\left. \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \end{pmatrix} \right\}.$$

If we arrange the elements in an egg-box diagram and put a star next to the idempotents we get the diagram in Figure 4.2. The distribution of idempotents in this example agrees with the general description given in Lemma 4.39.

Looking back at the statement of Theorem 4.24 observe that the exceptions to the rule in the theorem (i.e. those examples where idempotent rank and rank are not the same) are precisely the examples given in Lemma 4.39. All the other ideals of $P_{n,m}$ do have extremal idempotent generating sets.

Proposition 4.41. *Let $n, m \in \mathbb{N}$ with $n \geq 2$. Let $K'(n,r) = \{\alpha \in P_{n,m}$:*

$|\operatorname{im}\alpha \cap X_n| \leq r\}$ *for* $0 \leq r < n$. *Then we have*

$$\operatorname{idrank}(K'(n,r)) = \begin{cases} \operatorname{rank}(K'(n,r)) + 1 & \text{if } n = 2,\ r = 1 \text{ and } m \geq 1 \\ \operatorname{rank}(K'(n,r)) & \text{otherwise.} \end{cases}$$

Proof. Let $K(n,r) = \{\alpha \in T_n : |\operatorname{im}\alpha| \leq r\}$, $V_{n,r} = K(n,r)/K(n,r-1)$ and $Q_{n,r} = K'(n,r)/K'(n,r-1)$. There are a number of cases to consider depending on the value of r.

Case 1: $r = 0$. In this case $K'(n,r)$ is a band and as a consequence has an extremal idempotent generating set.

Case 2: $r \geq 2$. In this case, by Lemma 4.37, $V_{n,r}$ is an $\mathcal{L}$-class filling subsemigroup of $Q_{n,r}$ where, since $r \geq 2$, $V_{n,r}$ and $Q_{n,r}$ are both tall. The result now follows by Lemma 3.3.

Case 3: $r = 1$ *and* $n > 2$. Let D_1 denote the top $\mathcal{J}$-class of the semigroup $K'(n,1)$. The number of $\mathcal{L}$-classes in D_1 is equal to n, corresponding to the images

$$I_q = \{q, n+1, \ldots, n+m\},$$

where $1 \leq q \leq n$. Now consider the following n kernels:

$$K_j = \{\{j, j+1\}, X_{n+1} \setminus \{j, j+1\}, \{n+2\}, \ldots, \{n+m\}\}$$

for $1 \leq j \leq n-1$ and

$$K_n = \{\{n, 1\}, X_{n+1} \setminus \{n, 1\}, \{n+2\}, \ldots, \{n+m\}\}.$$

Note that for $1 \leq j \leq n-1$ the only images that form a transversal of the kernel K_j are I_j and I_{j+1}. Also, the only images that form a transversal of the kernel K_n are I_1 and I_n. In a similar way each image is a transversal of precisely 2 of the kernels $\{K_1, \ldots, K_n\}$. Let T be the natural rectangular 0-band homomorphic image of the principal factor $P_{n,1}$. The kernels listed correspond to a rectangular 0-band T_2 that is a tall (since it is square) $\mathcal{L}$-class filling subsemigroup of T. We have seen above that the number 1 appears exactly twice in every row and every column of the structure matrix of T_2. In other words the matrix has a symmetric distribution of idempotents. Also, it is easy to see that T_2 is a connected rectangular 0-band. Therefore, by Lemma 3.28, the rectangular 0-band T_2 has an extremal idempotent generating set which, by Lemma 3.3, implies that

T has an extremal idempotent generating set. It now follows from Lemma 4.16 that $K'(n,r)$ has an extremal idempotent generating set.

Case 4: $r = 1$ *and* $n = 2$. This is the one case where the idempotent rank and the rank are not the same. It is dealt with in Lemma 4.39. □

Trivial independence algebras where $\langle \varnothing \rangle \neq \varnothing$ and G is non-trivial

We show that when G is non-trivial the proper ideals of $\mathrm{End}(\mathcal{A})$ have extremal idempotent generating sets. We saw in the previous section that the number of $\mathcal{L}$-classes in $D_r \subseteq \mathrm{End}(\mathcal{A})$ is equal to $\binom{n}{r}$ for $1 \leq r \leq n$. Finding an expression for the number of $\mathcal{R}$-classes is slightly more tricky. We know from Lemma 4.37 that there are at least as many $\mathcal{R}$-classes as in the corresponding $\mathcal{D}$-class of $G \wr T_n$.

Definition 4.42. Let

$$\gamma(n,m,r,G) = \frac{1}{|G|^r} \sum_{k=0}^{n-r} m^k |G|^{n-k} \binom{n}{k} S(n-k,r)$$

where $|C| = m > 0$.

Lemma 4.43. *The formula* $\gamma(n,m,r,G)$ *gives the number of* $\mathcal{R}$*-classes of* $D_r \subseteq \mathrm{End}(\mathcal{A})$ *where* G *is non-trivial and* $1 \leq r \leq n-1$.

Proof. First we claim that

$$|D_r| = \binom{n}{r} r! \sum_{k=0}^{n-r} m^k |G|^{n-k} \binom{n}{k} S(n-k,r)$$

where $1 \leq r \leq n$. We count the number of distinct maps $\alpha \in \mathrm{Map}(B,\mathcal{A})$ that extend to endomorphisms $\overline{\alpha}$ with $\dim(\mathrm{im}\,\overline{\alpha}) = r$. There are $\binom{n}{r}$ choices for $\mathrm{im}\,\overline{\alpha}$, since $\mathcal{A}$ has $\binom{n}{r}$ r-dimensional subalgebras. Let $k = |\{b \in B : b\alpha \in C\}|$ noting that since $\dim(\mathrm{im}\,\overline{\alpha}) = r$ we have $0 \leq k \leq n-r$ (since $|B| = n$ and if more than $n-r$ elements mapped to constants then $\dim(\mathrm{im}\,\overline{\alpha}) < r$). There are $\binom{n}{k}$ possible choices for the set $\{b \in B : b\alpha \in C\}$ and there are m^k ways of assigning the images of the elements in $\{b \in B : b\alpha \in C\}$. The term $S(n-k,r)$ gives the number of ways of partitioning $\{b \in B : b\alpha \notin C\}$ into r kernel classes. There are $r!$ ways of assigning the images to these kernel classes and $|G|^{n-k}$ choices within the groups for each of the elements of the set $\{b \in B : b\alpha \notin C\}$ to map to. Summing as k runs from 0 up to $n-r$ gives the displayed formula above for $|D_r|$. Note that this formula does not hold for $C = \varnothing$. Let ρ and λ denote the number of $\mathcal{R}$- and $\mathcal{L}$-classes, respectively, of D_r. By Lemma 4.38, the group $\mathcal{H}$-classes of D_r are isomorphic to $G \wr S_r$. It follows that

$$\rho = \frac{|D_r|}{\lambda\, |G \wr S_r|} = \frac{|D_r|}{\binom{n}{r}\, |G \wr S_r|} = \frac{|D_r|}{\binom{n}{r}\, r!\, |G|^r} = \gamma(n, m, r, G).$$

□

To complete the chapter, we determine the idempotent rank of the proper ideals of the semigroup $\mathrm{End}(\mathcal{A})$, where $\mathcal{A}$ is a trivial independence algebra with constants, and G is non-trivial.

Theorem 4.44. *Let G be a finite group, let $m, n \in \mathbb{N}$, and let $\mathcal{A} = (X_n \times G) \cup C$ where $|C| = m$. Let*

$$K'(n, r) = \{\alpha \in \mathrm{End}(\mathcal{A}) : |\{i \in X_n : (i, g) = b\alpha \text{ for some } b \in B\}| \leq r\}$$

for $0 \leq r < n$. If $r = 0$ then

$$\mathrm{idrank}(K'(n, r)) = \mathrm{rank}(K'(n, r)) = m^n.$$

If G is non-trivial and $r \geq 1$ then

$$\mathrm{idrank}(K'(n, r)) = \mathrm{rank}(K'(n, r)) = \max(\binom{n}{r}, \gamma(n, m, r, G)).$$

If G is trivial and $r \geq 1$ then $\mathrm{End}(\mathcal{A}) \cong P_{n,m}$ and

$$\mathrm{idrank}(K'(n, r)) = \begin{cases} \max(\binom{n}{r}, \gamma(n, m, r, G)) + 1 & \text{if } n = 2,\ r = 1 \text{ and } m \geq 1 \\ \max(\binom{n}{r}, \gamma(n, m, r, G)) & \text{otherwise.} \end{cases}$$

Proof. The case when G is trivial is dealt with in Proposition 4.41. When G is non-trivial there are a number of cases to consider depending on the value of r.

Case 1: $r = 0$. In this case $K'(n, r)$ is a left zero semigroup, every $b \in B$ has C places it can map to giving m^n choices in total.

Case 2: $r = 1$. Let T be the rectangular 0-band homomorphic image of the principal factor corresponding to the unique maximal $\mathcal{J}$-class of $K'(n, r)$. Let $RB^0_{a,b}$ denote the $a \times b$ rectangular band with a zero adjoined. Then, by Lemma 4.37, $RB^0_{|G|^{n-1},n}$ is an $\mathcal{L}$-class filling subsemigroup of T. When G is non-trivial $|G|^{n-1} \geq n$ and $RB^0_{|G|^{n-1},n}$ is tall. The result now follows from Lemma 3.3 and the dimensions of the $\mathcal{D}$-class $D_r \subseteq \mathrm{End}(\mathcal{A})$.

Case 3: $r \geq 2$. Let T_1 and T_2 be the rectangular 0-band homomorphic images of the principal factors that correspond to the maximal $\mathcal{J}$-classes of $K(n,r)$ and $K'(n,r)$ respectively. By Theorem 4.34 the rectangular 0-band T_1 has an extremal idempotent generating set. By Lemma 4.37 the rectangular 0-band T_1 is an $\mathcal{L}$-class filling subsemigroup of T_2. Also, T_1 and T_2 are both tall. The result now follows from Lemma 3.3 and the dimensions of the $\mathcal{D}$-class $D_r \subseteq \text{End}(\mathcal{A})$. $\square$

Chapter 5

Vector spaces and non-trivial independence algebras

5.1 Non-trivial independence algebras

Let V be an n-dimensional vector space over the finite field F with $|F| = q = p^k$ where p is a prime. Let W be an m-dimensional subspace of V. As mentioned in the introduction to the previous chapter, a vector space is an independence algebra. It is a non-trivial independence algebra with a single constant (the zero vector). The basic properties of this algebra's endomorphism monoid, the general linear semigroup $\mathrm{GLS}(n, F)$, were given in Section 3.5. We saw that the ideals of the general linear semigroup have extremal idempotent generating sets in Theorem 3.39. More generally than this, the algebra $V[W]$ (using the construction described at the end of Section 4.3) is a finite independence algebra. It has dimension $n - m$ and its subalgebras are the subspaces of V that contain W and $\langle \varnothing \rangle = W$. Moreover, as a result of [14, Proposition 8.1], every finite non-trivial independence algebra with constants is equivalent to one constructed from a vector space and a subspace in this way.

In the following section we will describe the structure of the semigroup $\mathrm{End}(V[W])$. In particular, a detailed description of the distribution of the idempotents in the semigroup will be provided. This will then be used to determine the rank and idempotent rank of the ideals of the semigroup.

5.2 Non-trivial independence algebras with constants

Let $\mathrm{Mat}_{a\times b}(F)$ denote the set of all $a \times b$ matrices over the field F. The endomorphism monoid $\mathrm{End}(V[W])$ consists of all linear transformations of V that fix W elementwise. Concretely we may represent $\mathrm{End}(V[W])$ as the subsemigroup of $\mathrm{GLS}(n, F)$ defined by:

$$\left\{ \begin{pmatrix} I_m & 0 \\ A & X \end{pmatrix} : A \in \mathrm{Mat}_{(n-m)\times m}(F),\ X \in \mathrm{GLS}(n-m, F) \right\} \leq \mathrm{GLS}(n, F)$$

where I_m denotes the $m \times m$ identity matrix. Multiplying two of these matrices together gives:

$$\begin{pmatrix} I_m & 0 \\ A & X \end{pmatrix} \begin{pmatrix} I_m & 0 \\ B & Y \end{pmatrix} = \begin{pmatrix} I_m & 0 \\ A + XB & XY \end{pmatrix}.$$

Therefore $\mathrm{End}(V[W])$ is isomorphic to the semigroup of pairs

$$\mathrm{Mat}_{(n-m)\times m}(F) \times \mathrm{Mat}_{(n-m)\times(n-m)}(F)$$

with multiplication:

$$(A, X)(B, Y) = (A + XB, XY).$$

For what remains of this section we will think of $\mathrm{End}(V[W])$ as this semigroup of pairs. An important subsemigroup of $\mathrm{End}(V[W])$ is

$$N = \{(0, X) : X \in \mathrm{GLS}(n - m, F)\} \leq \mathrm{End}(V[W]).$$

Let $U = F^{n-m}$ and observe that $N \cong \mathrm{End}(U) \cong \mathrm{GLS}(n - m, F)$. In much the same way as T_n dictated the structure of $G \wr T_n$ (see Section 4.7), the semigroup $\mathrm{End}(U)$ will determine the structure of $\mathrm{End}(V[W])$.

We start by describing Green's relations in the semigroup. Given an element (A, X) in $\mathrm{End}(V[W])$ we view X as belonging to the general linear semigroup $\mathrm{GLS}(n - m, F)$ and we write $\dim X$ to mean the dimension of the image of the linear transformation X in $\mathrm{GLS}(n - m, F)$. Thus $\dim X$ lies somewhere between 0 and $n - m$.

Lemma 5.1. *Let* $(A, X), (B, Y) \in \mathrm{End}(V[W])$. *Then*

(i) $(A, X)\mathcal{L}(B, Y)$ *in* $\mathrm{End}(V[W])$ *if and only if* $X\mathcal{L}Y$ *in* $\mathrm{End}(U)$.

(ii) $(A, X)\mathcal{R}(B, Y)$ *in* $\mathrm{End}(V[W])$ *if and only if* $X\mathcal{R}Y$ *in* $\mathrm{End}(U)$ *and there exists some* $Z \in \mathrm{Mat}_{(n-m)\times m}(F)$ *such that* $YZ = A - B$.

(iii) $(A, X)\mathcal{D}(B, Y)$ *if and only if* $\dim(\mathrm{im}\, X) = \dim(\mathrm{im}\, Y)$.

Proof. (i) From the definition of multiplication in the semigroup it is seen that (A, X) and (B, Y) are $\mathcal{L}$-related in $\mathrm{End}(V[W])$ if and only if X and Y are $\mathcal{L}$-related in $\mathrm{End}(U)$.

(ii) The elements (A, X) and (B, Y) are $\mathcal{R}$-related in $\mathrm{End}(V[W])$ if and only if there exist (D_1, Q_1) and (D_2, Q_2) in $\mathrm{End}(V[W])$ such that

$$(A + XD_1, XQ_1) = (A, X)(D_1, Q_1) = (B, Y)$$

and

$$(B + YD_2, YQ_2) = (B, Y)(D_2, Q_2) = (A, X).$$

This is equivalent to saying that X and Y are $\mathcal{R}$-related in $\mathrm{End}(U)$, $XD_1 = B - A$ and $YD_2 = A - B$. However, if $YD_2 = A - B$ then $XQ_1D_2 = A - B$ and it follows that $X(-I)Q_1D_2 = B - A$. It follows that if there exists a matrix D_1

satisfying $XD_1 = B - A$ then there exists a matrix D_2 such that $YD_2 = A - B$. The converse also holds and so the two conditions are equivalent and one of them may be omitted.

(iii) Follows from Proposition 4.15. □

Now we will find where the idempotents are by identifying the group $\mathcal{H}$-classes.

Lemma 5.2. *The $\mathcal{H}$-class $H_{(A,X)}$ in* End$(V[W])$ *is a group if and only if H_X is a group in* End(U).

Proof. The $\mathcal{H}$-class $H_{(A,X)}$ is a group if and only if (A, X) is $\mathcal{H}$-related to $(A, X)^2 = (A + XA, X^2)$. This is equivalent, by Lemma 5.1, to saying that X is $\mathcal{H}$-related to X^2 in End(U) and there exists a $Z \in \mathrm{Mat}_{(n-m)\times m}(F)$ such that $XZ = (A + XA) - A = XA$. This equation is satisfied by setting $Z = A$ and the result follows. □

The semigroup End$(V[W])$ has $n - m + 1$ $\mathcal{D}$-classes given by

$$D_r = \{(A, X) \in \mathrm{End}(V[W]) : \dim(\mathrm{im}\, X) = r\}$$

where $0 \leq r \leq n - m$. Also, from Lemma 5.1(i), it follows that the $\mathcal{L}$-classes of D_r are in one-one correspondence with the r-dimensional subspaces of End(U).

Now consider the $\mathcal{R}$-classes. Our first step will be to partition the set of $\mathcal{R}$-classes in D_r. Let H be the group of all $(n - m) \times m$ matrices over the field F under addition. This group is just the direct power of $m(n - m)$ copies of the additive group of the field F. The group H acts on the $\mathcal{R}$-classes of D_r in the following way. Let $A \in H$ and let R be an $\mathcal{R}$-class of D_r. Then define

$$A \cdot R = (A, I)R = \{(A, I)(B, Y) : (B, Y) \in R\}$$

where I denotes the $(n-m)\times(n-m)$ identity matrix. The element $(A, I)(B, Y) = (A+B, Y)$ is $\mathcal{D}$-related to (B, Y), by Lemma 5.1, and so by Green's lemma $(A, I)R$ is an $\mathcal{R}$-class of End$(V[W])$ in D_r. Since $0 \cdot R = (0, I)R = R$ and

$$(AB) \cdot R = ((A, I)(B, I))R = (A, I)((B, I)R) = A \cdot (B \cdot R)$$

this is a group action. Denote the orbit of the $\mathcal{R}$-class $R_{(A,X)}$ by $\mathcal{O}_{(A,X)}$.

As in the previous section, we now prove that each of the orbits has the same size, that the $\mathcal{R}$-classes of D_r in N form a transversal of the orbits, and that any two $\mathcal{R}$-classes in the same orbit look the same in terms of the position of idempotents (i.e. that the map $f : R \to (A, I)R$ defined by $x \mapsto (A, I)x$ sends group $\mathcal{H}$-classes to group $\mathcal{H}$-classes).

Lemma 5.3. *The $\mathcal{R}$-classes of $N \cap D_r$ form a transversal of the orbits of the $\mathcal{R}$-classes of D_r.*

Proof. Let $(A, X) \in D_r$. Then $(-A, I)(A, X) = (0, X) \in N$ and so the set of $\mathcal{R}$-classes of $N \cap D_r$ intersects every orbit at least once.

To see that no orbit is intersected more than once let $\mathcal{O}_{(0,X)} = \mathcal{O}_{(0,Y)}$. Then $(D, I)R_{(0,X)} = R_{(0,Y)}$ for some (D, I) where $D \in H$. Therefore $(D, X) \in R_{(0,Y)}$ giving $X\mathcal{R}Y$ in $\mathrm{End}(U)$ (by Lemma 5.1) and it follows that $R_{(0,X)} = R_{(0,Y)}$. □

Lemma 5.4. *Let $(A, X), (B, I) \in \mathrm{End}(V[W])$. Then $H_{(A,X)}$ is a group if and only if $H_{(B,I)(A,X)}$ is a group.*

Proof. The $\mathcal{H}$-class $H_{(A,X)}$ is a group, by Lemma 5.2, if and only if H_X is a group in $\mathrm{End}(U)$ which is true if and only if $H_{(B+A,X)} = H_{(B,I)(A,X)}$ is a group in $\mathrm{End}(V[W])$. □

Now we prove that the orbits all have the same size. Before we do this, however, we need to make another definition. For each $(0, Q) \in N$ define

$$\mathcal{G}_{(0,Q)} = \{A \in \mathrm{Mat}_{(n-m)\times m}(F) : (A, X) \in R_{(0,Q)} \text{ for some } X \in \mathrm{GLS}(n-m, F)\}.$$

If $(A, X), (B, Y) \in R_{(0,Q)}$ then by Lemma 5.1 there exist $Z_1, Z_2 \in \mathrm{Mat}_{(n-m)\times m}(F)$ such that $QZ_1 = A$ and $QZ_2 = B$ so that $Q(Z_1 + Z_2) = A + B$ which, along with the fact that $X\mathcal{R}Q$ in $\mathrm{End}(U)$, gives $(A + B, X) \in R_{(0,Q)}$. In a similar way one can prove that $(A + B, Y) \in R_{(0,Q)}$. It follows from this that $\mathcal{G}_{(0,Q)}$ is a subgroup of the group H of all $(n - m) \times m$ matrices over F under addition. The next lemma justifies the definition of the group $\mathcal{G}_{(0,X)}$.

Lemma 5.5. *Let $X \in \mathrm{GLS}(n - m, F)$. Then the number of $\mathcal{R}$-classes in $\mathcal{O}_{(0,X)}$ is equal to the index of $\mathcal{G}_{(0,X)}$ in H.*

Proof. For each $\mathcal{R}$-class R in $\mathcal{O}_{(0,X)}$ let

$$L(R) = \{A \in \mathrm{Mat}_{(n-m)\times m}(F) : (A, X) \in R \text{ for some } X \in \mathrm{GLS}(n-m, F)\}.$$

Since $(A, I)(0, X) = (A, X)$ it follows that $\bigcup_{R\in\mathcal{O}_{(0,X)}} L(R) = H$. From this, and from the definition of the action, it follows that $\{L(R) : R \in \mathcal{O}_{(0,X)}\}$ is precisely the set of all cosets of the subgroup $\mathcal{G}_{(0,X)}$ of the additive group H. Moreover, it follows from the definition of the action that if $R_1, R_2 \in \mathcal{O}_{(0,X)}$ and

$L(R_1) = L(R_2)$ then $R_1 = R_2$. Thus, there is a one-one correspondence between $\mathcal{R}$-classes in $\mathcal{O}_{(0,X)}$ and cosets of $\mathcal{G}_{(0,X)}$ in H. □

Therefore, proving $|\mathcal{O}_{(0,X)}| = |\mathcal{O}_{(0,Y)}|$ is equivalent to proving that $|\mathcal{G}_{(0,X)}| = |\mathcal{G}_{(0,Y)}|$.

Lemma 5.6. *Let $(0, X)$ and $(0, Y)$ be in D_r. Then $\mathcal{G}_{(0,X)}$ and $\mathcal{G}_{(0,Y)}$ are isomorphic groups.*

Proof. We can suppose without loss of generality that X, Y are $\mathcal{L}$-related in $\mathrm{End}(U)$. This is because X and Y are $\mathcal{D}$-related in $\mathrm{End}(U)$ and so there is an X' such that $X\mathcal{R}X'\mathcal{L}Y$ satisfying $\mathcal{G}_{(0,X)} = \mathcal{G}_{(0,X')}$. Let $U_1, U_2 \in \mathrm{End}(U)$ so that

$$U_1X = Y, \quad U_2Y = X.$$

By Lemma 5.1

$$\begin{aligned}
\mathcal{G}_{(0,X)} &= \{A \in \mathrm{Mat}_{(n-m)\times m}(F) : (A, K) \in R_{(0,X)} \text{ for some } K \in \mathrm{GLS}(n-m, F)\} \\
&= \{A \in \mathrm{Mat}_{(n-m)\times m}(F) : (A, K)\mathcal{R}(0, X) \text{ for some } K \in \mathrm{GLS}(n-m, F)\} \\
&= \{A \in \mathrm{Mat}_{(n-m)\times m}(F) : K\mathcal{R}X \text{ in } \mathrm{End}(U) \text{ and} \\
&\qquad\qquad \text{there exists } Z \in \mathrm{Mat}_{(n-m)\times m}(F) : XZ = A \\
&\qquad\qquad\qquad \text{for some } K \in \mathrm{GLS}(n-m, F)\} \\
&= \{A \in \mathrm{Mat}_{(n-m)\times m}(F) : A = XZ \text{ for some } Z \in \mathrm{Mat}_{(n-m)\times m}\}
\end{aligned}$$

and similarly

$$\mathcal{G}_{(0,Y)} = \{A \in \mathrm{Mat}_{(n-m)\times m}(F) : A = YZ \text{ for some } Z \in \mathrm{Mat}_{(n-m)\times m}\}.$$

Define $\phi : \mathcal{G}_{(0,X)} \to \mathcal{G}_{(0,Y)}$ by $A\phi = B$ where $A = XZ$ and $B = YZ$. This map is well defined since if $A = XZ_1 = XZ_2$ then $YZ_1 = U_1XZ_1 = U_1XZ_2 = YZ_2$. Now define $\phi' : \mathcal{G}_{(0,Y)} \to \mathcal{G}_{(0,X)}$ by $B\phi' = A$ where $B = YZ$ and $A = XZ$. This map is well defined by the same argument as above. Now we have

$$XZ\phi\phi' = YZ\phi' = XZ$$

and it follows that ϕ is a bijection. Also, ϕ is a homomorphism since:

$$(XZ_1+XZ_2)\phi = (X(Z_1+Z_2))\phi = Y(Z_1+Z_2) = YZ_1+YZ_2 = (XZ_1)\phi+(XZ_2)\phi.$$

□

Now we compute the size of the group $\mathcal{G}_{(0,Q)}$. We will then use this number to find the number of $\mathcal{R}$-classes in D_r. The field F is a vector space over $P = \{1, 1+1, \ldots, \underbrace{1+\ldots+1}_{p}\}$, and $P \cong \mathbb{Z}_p$. The additive abelian group of the field F is isomorphic to $\mathbb{Z}_p^k$ and so $H \cong \mathbb{Z}_p^{km(n-m)}$, the direct power of $m(n-m)$ copies of $\mathbb{Z}_p^k$.

Lemma 5.7. *If* $(A, X) \in D_r$ *then* $|\mathcal{O}_{(A,X)}| = q^{m(n-m-r)}$.

Proof. From the proof of Lemma 5.6 it follows that

$$\mathcal{G}_{(0,Q)} = \{A \in \mathrm{Mat}_{(n-m)\times m}(F) : A = QZ \text{ for some } Z \in \mathrm{Mat}_{(n-m)\times m}(F)\}.$$

Let Q be the $(n-m) \times (n-m)$ matrix

$$Q = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

which has an $r \times r$ identity matrix in the top left corner and zeros everywhere else. Then

$$\mathcal{G}_{(0,Q)} = \left\{ \begin{pmatrix} A \\ 0 \end{pmatrix} : A \in \mathrm{Mat}_{r\times m}(F) \right\} \subseteq \mathrm{Mat}_{(n-m)\times m}(F).$$

This group is isomorphic to $\mathbb{Z}_p^{krm}$. It follows that:

$$|\mathcal{O}_{(0,Q)}| = [H : \mathcal{G}_{(0,Q)}] = \frac{|H|}{|\mathcal{G}_{(0,Q)}|} = \frac{p^{km(n-m)}}{p^{krm}} = q^{m(n-m-r)}.$$

Since, by Lemma 5.6, all the groups $\mathcal{G}_{(0,X)}$ are isomorphic it follows that for every matrix $P \in \mathrm{GLS}(n-m, F)$ we have $|\mathcal{O}_{(0,P)}| = q^{m(n-m-r)}$. □

Combining these lemmas together gives the following result.

Proposition 5.8. *Let* V *be a vector space with dimension* n *over the finite field* F, *with* $|F| = q$, *let* W *be an* m*-dimensional subspace of* V, *and let* $U = F^{n-m}$. *Let*

$$I(n, q, r) = \{(A, X) \in \mathrm{End}(V[W]) : \dim(\operatorname{im} X) \leq r\}$$

and

$$I'(n, q, r) = \{X \in \mathrm{End}(U) : \dim(\operatorname{im} X) \leq r\}$$

for $0 \leq r \leq n-m$. *Also, let* $V_{n,r} = I(n,q,r)/I(n,q,r-1)$, $Q_{n,r} =$

$I'(n,q,r)/I'(n,q,r-1)$ *and*

$$D_r = \{(A, X) \in \mathrm{End}(V[W]) : \dim(\mathrm{im}\, X) = r\}.$$

Then we have:

(i) The number of $\mathcal{L}$-classes in D_r is $\begin{bmatrix} n-m \\ r \end{bmatrix}_q$.

(ii) The number of $\mathcal{R}$-classes in D_r is $q^{m(n-m-r)}\begin{bmatrix} n-m \\ r \end{bmatrix}_q$.

(iii) Let T_1 and T_2 be the 0-rectangular band homomorphic images of $V_{n,r}$ and $Q_{n,r}$ respectively. Then $T_1 \cong {L^0}_{q^{m(n-m-r)}} \times_0 T_2$ where ${L^0}_k$ denotes the k-element left zero semigroup with a zero adjoined.

Proof. (i) By Lemma 5.1 the $\mathcal{L}$-classes are in one-one correspondence with the r-dimensional subspaces of the vector space U. Since $\dim(U) = n - m$ this number is given by the expression in the proposition.

(ii) Since N is isomorphic to U, the number of $\mathcal{R}$-classes in $D_r \cap N$ is equal to $\begin{bmatrix} n-m \\ r \end{bmatrix}_q$. The result now follows from Lemma 5.3 and Lemma 5.7.

(iii) Follows from Lemmas 5.3, 5.4 and 5.7. □

Example 5.9. Let $V = \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$, a 4-dimensional vector space over the field with 2 elements, and let $W = \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \{0\} \oplus \{0\}$, a 2-dimensional subspace of V. In this example $U \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2$. The independence algebra $V[W]$ has dimension $\dim(V) - \dim(W) = 2$. The endomorphism monoid of the algebra is isomorphic to

$$\mathrm{End}(V[W]) = \left\{ \begin{pmatrix} I_2 & 0 \\ A & B \end{pmatrix} : A, B \in \mathrm{GLS}(2, \mathbb{Z}_2) \right\} \leq \mathrm{GLS}(4, \mathbb{Z}_2).$$

The semigroup $\mathrm{End}(V[W])$ has 2^8 elements, it has 3 $\mathcal{D}$-classes corresponding to maps with image dimension 2, 1 and 0. Figure 5.1 shows how the structure of $\mathrm{End}(V[W])$ and of $\mathrm{End}(U)$ are related to one another. The middle $\mathcal{D}$-class of $\mathrm{End}(V[W])$ is built from 4 copies of the middle $\mathcal{D}$-class of $\mathrm{End}(U)$, stacked on top of one another. The bottom $\mathcal{D}$-class of $\mathrm{End}(V[W])$ is built from 16 copies of the bottom $\mathcal{D}$-class of $\mathrm{End}(U)$, stacked on top of one another. In the language of Section 3.2, the rectangular 0-band homomorphic images of the principal factors of $\mathrm{End}(V[W])$ are isomorphic to direct products, with amalgamated zero, of certain left zero semigroups and the rectangular 0-band homomorphic images of

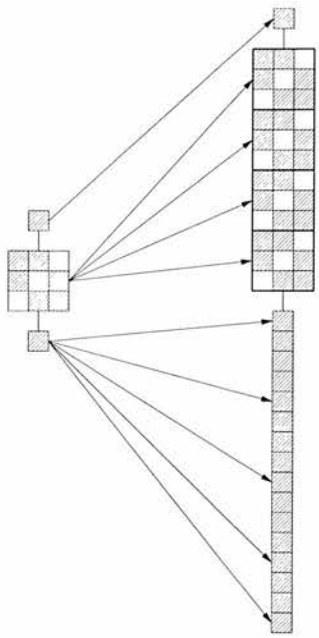


Figure 5.1: Comparing the egg-box pictures of the semigroup End(U) where $U = \mathbb{Z}_2 \oplus \mathbb{Z}_2$, on the left, and the semigroup End($V[W]$) where $V = \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$ and $W = \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \{0\} \oplus \{0\}$, on the right.

the principal factors of End(U). Notice how similar this figure is to Figure 4.1 which showed the relationship between the structure of T_n and that of End($G \wr T_n$).

In the next example we compute all the elements of a particular $\mathcal{D}$-class and then describe the orbits of the $\mathcal{R}$-classes under the action of the group H.

Example 5.10. Let $V = \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$ and $W = \mathbb{Z}_2 \oplus \{0\} \oplus \{0\}$. The independence algebra $V[W]$ has dimension 2 and the endomorphism monoid End($V[W]$) has three $\mathcal{D}$-classes. The middle $\mathcal{D}$-class is illustrated in Figure 5.2 with the elements arranged into $\mathcal{R}$-, $\mathcal{L}$- and $\mathcal{H}$-classes. Here $n = 3$, $m = 1$ and the group H is the additive group of 2×1 matrices over $\mathbb{Z}_2$. This group is isomorphic to $\mathbb{Z}_2 \oplus \mathbb{Z}_2$. The orbits of the set $\{R_1, \dots, R_6\}$ of $\mathcal{R}$-classes under the action of the group H are given by:

$$\mathcal{O}_{[\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}]} = \mathcal{O}_{[\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}]} = \{R_1, R_6\},$$

$$\mathcal{O}_{[\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix}]} = \mathcal{O}_{[\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix}]} = \{R_2, R_4\},$$

$$\mathcal{O}_{[\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}0&1\\0&1\end{pmatrix}]} = \mathcal{O}_{[\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}0&1\\0&1\end{pmatrix}]} = \{R_3, R_5\}.$$

R_1	$\begin{pmatrix}1&0&0\\0&0&0\\0&0&1\end{pmatrix}$ $\begin{pmatrix}1&0&0\\0&0&0\\1&0&1\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&0&0\\0&1&0\end{pmatrix}$ $\begin{pmatrix}1&0&0\\0&0&0\\1&1&0\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&0&0\\0&1&1\end{pmatrix}$ $\begin{pmatrix}1&0&0\\0&0&0\\1&1&1\end{pmatrix}$
R_2	$\begin{pmatrix}1&0&0\\0&0&1\\0&0&0\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&0&1\\0&0&0\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&1&0\\0&0&0\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&1&1\\0&0&0\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&1&1\\0&0&0\end{pmatrix}$
R_3	$\begin{pmatrix}1&0&0\\0&0&1\\0&0&1\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&0&1\\1&0&1\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&1&0\\0&1&0\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&1&0\\1&1&0\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&1&1\\0&1&1\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&1&1\\1&1&1\end{pmatrix}$
R_4	$\begin{pmatrix}1&0&0\\0&0&1\\1&0&0\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&0&1\\1&0&0\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&1&0\\1&0&0\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&1&0\\1&0&0\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&1&1\\1&0&0\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&1&1\\1&0&0\end{pmatrix}$
R_5	$\begin{pmatrix}1&0&0\\0&0&1\\1&0&1\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&0&1\\0&0&1\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&1&0\\1&1&0\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&1&0\\0&1&0\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&1&1\\1&1&1\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&1&1\\0&1&1\end{pmatrix}$
R_6	$\begin{pmatrix}1&0&0\\1&0&0\\0&0&1\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&0&0\\1&0&1\end{pmatrix}$	$\begin{pmatrix}1&0&0\\1&0&0\\0&1&0\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&0&0\\1&1&0\end{pmatrix}$	$\begin{pmatrix}1&0&0\\1&0&0\\0&1&1\end{pmatrix}$ $\begin{pmatrix}1&0&0\\1&0&0\\1&1&1\end{pmatrix}$

Figure 5.2: The elements of the middle $\mathcal{D}$-class of the monoid $\mathrm{End}(V[W])$ where $V = \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$ and $W = \mathbb{Z}_2 \oplus \{0\} \oplus \{0\}$.

The set $N \cap J = \{R_1, R_2, R_3\}$ is indeed a transversal of the orbits above, as Lemma 5.3 says it must be. Using this example we can work through the steps of the proof of Lemma 5.6. Let

$$(0, X) = \left[\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\right], \quad (0, Y) = \left[\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}\right]$$

noting that $X\mathcal{L}Y$ in $\mathrm{End}(\mathbb{Z}_2 \oplus \mathbb{Z}_2)$ since with

$$U_1 = \begin{pmatrix}1 & 1\\1 & 0\end{pmatrix}, \quad U_2 = \begin{pmatrix}0 & 1\\1 & 1\end{pmatrix}$$

we have

$$U_1X = \begin{pmatrix}1 & 1\\1 & 0\end{pmatrix}\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix} = \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix} = Y$$

and

$$U_2X = \begin{pmatrix}0 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}0 & 1\\0 & 0\end{pmatrix} = \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix} = X.$$

Now

$$\mathcal{G}_{(0,X)} = \mathcal{G}_{\left[\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\right]} = \left\{\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$$

and

$$\mathcal{G}_{(0,Y)} = \mathcal{G}_{\left[\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}\right]} = \left\{\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right\}.$$

The map $\phi : \mathcal{G}_{(0,X)} \to \mathcal{G}_{(0,Y)}$ is defined by $A\phi = B$ where $A = XZ$ and $B = YZ$ for some $Z \in \mathrm{Mat}_{2\times 1}(\mathbb{Z}_2)$. Therefore

$$\begin{pmatrix}0\\0\end{pmatrix}\phi = \begin{pmatrix}0\\0\end{pmatrix}, \quad \begin{pmatrix}0\\1\end{pmatrix}\phi = \begin{pmatrix}1\\0\end{pmatrix}.$$

The map is an isomorphism from the additive group of matrices $\mathcal{G}_{(0,X)}$ to the additive group of matrices $\mathcal{G}_{(0,Y)}$. For every $X \in \mathrm{GLS}(2, \mathbb{Z}_2)$ the group $\mathcal{G}_{(0,X)}$ has index 2 in H which, as Lemma 5.5 says it must do, equals the size of each of the orbits of the $\mathcal{R}$-classes.

We now determine the idempotent rank of the ideals of $\mathrm{End}(V[W])$.

Theorem 5.11. *Let V be a vector space with dimension n over the finite field*

F, with $|F| = q$ and let W be an m-dimensional subspace of V. Let

$$I(n,q,r) = \{(A,X) \in \mathrm{End}(V[W]) : \dim(\mathrm{im}\, X) \le r\}$$

for $0 \le r \le n-m$. Then we have

$$\mathrm{idrank}(I(n,q,r)) = \mathrm{rank}(I(n,q,r)) = q^{m(n-m-r)} \begin{bmatrix} n-m \\ r \end{bmatrix}_q.$$

Proof. By Lemma 4.16 we know that the rank and idempotent rank of $I(n,q,r)$ equal the rank and idempotent rank of $V_{n,r}$. Now the result follows from the fact that $Q_{n,r}$ is square, Lemma 3.6, Proposition 5.8(iii) and Theorem 3.39. □

5.3 Non-trivial independence algebras without constants

Affine independence algebras

We begin with a brief discussion of affine groups and semigroups (see [7] for more details). The *n-dimensional affine group* over the field F is

$$\mathrm{Aff}_n(F) = \left\{ \begin{pmatrix} A & 0 \\ v & 1 \end{pmatrix} : A \in \mathrm{GL}(n,F), v \in F^n \right\} \le \mathrm{GL}(n+1,F).$$

Actually, the affine group is often given as the set of transposes of the above matrices, but these two groups are isomorphic (under ϕ defined by $A \mapsto (A^T)^{-1}$). We use the matrices above because we are viewing our matrices as acting on the vector space V on the right. Identify $x \in F^n$ with the vector $(x,1) \in F^{n+1}$ so that

$$(x,1)\begin{pmatrix} A & 0 \\ v & 1 \end{pmatrix} = (xA+v,1)$$

which gives an action of $\mathrm{Aff}_n(F)$ on F^n. Transformations of F^n with the form $x \mapsto xA+v$ are called *affine transformations*. The vector space F^n can be viewed as the *translational subgroup* of $\mathrm{Aff}_n(F)$

$$\mathrm{Trans}_n(F) = \left\{ \begin{pmatrix} I & 0 \\ v & 1 \end{pmatrix} : v \in F^n \right\} \le \mathrm{Aff}_n(F)$$

since

$$\begin{pmatrix} I & 0 \\ v & 1 \end{pmatrix}\begin{pmatrix} I & 0 \\ w & 1 \end{pmatrix} = \begin{pmatrix} I & 0 \\ v+w & 1 \end{pmatrix}$$

the group $\mathrm{Trans}_n(F)$ is isomorphic to the additive group of the vector space F^n. Another important subgroup is

$$\left\{ \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} : A \in \mathrm{GL}(n, F) \right\} \leq \mathrm{Aff}_n(F)$$

which may be identified with the group $\mathrm{GL}(n, F)$. It is now straightforward to verify that $\mathrm{Trans}_n(F)$ is a normal subgroup of $\mathrm{Aff}_n(F)$, that

$$\mathrm{Aff}_n(F) = \mathrm{GL}(n, F)\mathrm{Trans}_n(F) = \{gt : g \in \mathrm{GL}(n, F),\ t \in \mathrm{Trans}_n(F)\}$$

and finally that $\mathrm{GL}(n, F) \cap \mathrm{Trans}_n(F) = \{0\}$ (this is because the every linear transformation fixes the zero vector while the only translation that does so is translation by the zero vector itself). These three conditions holding imply that $\mathrm{Aff}_n(F)$ is a semidirect product of $\mathrm{Trans}_n(F)$ by $\mathrm{GL}(n, F)$ where the action of $\mathrm{GL}(n, F)$ on $\mathrm{Trans}_n(F)$ is given by

$$\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} I & 0 \\ v & 1 \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I & 0 \\ vA & 1 \end{pmatrix}.$$

Related to the group $\mathrm{Aff}_n(F)$ is the *n-dimensional affine semigroup*

$$\mathrm{AffSg}_n(F) = \left\{ \begin{pmatrix} A & 0 \\ v & 1 \end{pmatrix} : A \in \mathrm{GLS}(n, F), v \in F^n \right\} \leq \mathrm{GLS}(n+1, F).$$

In a similar way, this semigroup may be viewed as a semidirect product. Indeed, since

$$\begin{pmatrix} A & 0 \\ \alpha & 1 \end{pmatrix} \begin{pmatrix} B & 0 \\ \beta & 1 \end{pmatrix} = \begin{pmatrix} AB & 0 \\ \alpha B + \beta & 1 \end{pmatrix}$$

it follows that the semigroup $\mathrm{AffSg}_n(F)$ has a description as the semigroup of pairs $V \times \mathrm{End}(V)$ with multiplication:

$$(\alpha, A)(\beta, B) = (\alpha B + \beta, AB).$$

We will use $\mathrm{End}(\mathrm{Aff}(V))$ to denote this semigroup of pairs.

The connection between affine semigroups and non-trivial independence algebras without constants is given by [14, Example 3.2]. Let V be a vector space over a finite field F where $|F| \geq 3$. For each $c \in F$ with $c \neq 0, 1$, define a binary operator μ_c on V by $\mu_c(x, y) = x + c(y - x)$. The set V, with operations μ_c, is an independence algebra whose subalgebras are the affine subspaces of V (i.e. cosets

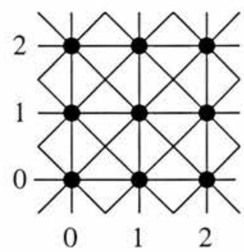


Figure 5.3: The affine subspaces of $V = \mathbb{Z}_3 \oplus \mathbb{Z}_3$.

of subspaces). If $|F| = 2$ then the construction must be adapted by adding the ternary operator $\alpha(x, y, z) = x + y + z$. In either case the algebra is written as $\text{Aff}(V)$.

The algebra $\text{Aff}(V)$ has dimension $n + 1$ (not n) and as a result has exactly this many $\mathcal{D}$-classes. The endomorphism monoid of $\text{Aff}(V)$ is isomorphic to the n-dimensional affine semigroup described above.

Example 5.12. Let $V = \mathbb{Z}_3 \oplus \mathbb{Z}_3$, where $\mathbb{Z}_3 = \{0, 1, 2\}$, and consider the algebra $\text{Aff}(V)$. It has a single binary operation defined by:

$$\mu_2(x, y) = x + 2(y - x) = 2y - x.$$

The independence algebra $\text{Aff}(V)$ has dimension 3. The 1-dimensional subalgebras are the singletons (which are cosets of the trivial subspace of V). The one dimensional subspaces of the vector space V are $\{(0, 0), (1, 0), (2, 0)\}$, $\{(0, 0), (0, 1), (0, 2)\}$, $\{(0, 0), (1, 1), (2, 2)\}$ and $\{(0, 0), (1, 2), (2, 1)\}$. Taking all translates (cosets) of these subspaces gives all the 2-dimensional subalgebras of $\text{Aff}(V)$. These affine subspaces correspond to horizontal, vertical and diagonal straight lines in $\mathbb{Z}_3 \oplus \mathbb{Z}_3$. Any two distinct points in $\mathbb{Z}_3 \oplus \mathbb{Z}_3$ extend uniquely to one of these lines and this is the affine subspace generated by these two points (see Figure 5.3). For example $\{(2, 0), (0, 2)\}$ extends to the line $\{(1, 1), (2, 0), (0, 2)\}$. Any three non-collinear points generate the whole algebra $\text{Aff}(V)$ and these 3-sets are the bases of the algebra.

As in the previous sections, our first step is to determine Green's relations in the semigroup $\text{End}(\text{Aff}(V))$.

Lemma 5.13. *Let* $(\alpha, A), (\beta, B) \in \text{End}(\text{Aff}(V))$. *Then*

(i) $(\alpha, A)\mathcal{L}(\beta, B)$ *in* $\text{End}(\text{Aff}(V))$ *if and only if* $A\mathcal{L}B$ *in* $\text{End}(V)$ *and* $\alpha - \beta \in \text{im}\, A = \text{im}\, B$.

(ii) $(\alpha, A)\mathcal{R}(\beta, B)$ *in* $\mathrm{End}(\mathrm{Aff}(V))$ *if and only if* $A\mathcal{R}B$ *in* $\mathrm{End}(V)$.

(iii) $(\alpha, A)\mathcal{D}(\beta, B)$ *in* $\mathrm{End}(\mathrm{Aff}(V))$ *if and only if* $\dim(\operatorname{im} A) = \dim(\operatorname{im} B)$.

Proof. (i) The elements (α, A) and (β, B) are $\mathcal{L}$-related in $\mathrm{End}(\mathrm{Aff}(V))$ if and only if there exist (γ, C) and (δ, D) in $\mathrm{End}(V[W])$ such that

$$(\gamma A + \alpha, CA) = (\gamma, C)(\alpha, A) = (\beta, B)$$

and

$$(\delta B + \beta, DB) = (\delta, D)(\beta, B) = (\alpha, A).$$

This is equivalent to saying that A and B are $\mathcal{L}$-related in $\mathrm{End}(V)$, $\gamma A = \beta - \alpha$ and $\delta B = \alpha - \beta$. Since $\operatorname{im} A = \operatorname{im} B$ is closed under addition it follows that $\alpha - \beta \in \operatorname{im} A = \operatorname{im} B$ if and only if $\beta - \alpha \in \operatorname{im} A = \operatorname{im} B$. We conclude that the last of these conditions is redundant and we can exclude it.

(ii) The elements (α, A) and (β, B) are $\mathcal{R}$-related in $\mathrm{End}(\mathrm{Aff}(V))$ if and only if there exist (γ, C) and (δ, D) in $\mathrm{End}(V[W])$ such that

$$(\alpha C + \gamma, AC) = (\alpha, A)(\gamma, C) = (\beta, B)$$

and

$$(\beta D + \delta, DB) = (\beta, B)(\delta, D) = (\alpha, A).$$

This is equivalent to saying that A and B are $\mathcal{R}$-related in $\mathrm{End}(V)$, $\gamma = \beta - \alpha C$ and $\delta = \alpha - \beta D$. We can just define γ and δ to equal these vectors.

(iii) Follows from Proposition 4.15. □

As already mentioned, the algebra $\mathrm{Aff}(V)$ has dimension $n + 1$ and so the $\mathcal{D}$-classes are given by:

$$D_r = \{(\alpha, A) \in \mathrm{End}(\mathrm{Aff}(V)) : \dim(\operatorname{im} A) = r\}$$

for $0 \leq r \leq n$. The $\mathcal{R}$-classes are in one-one correspondence with the $\mathcal{R}$-classes of $\mathrm{End}(V)$. Note that one big difference between the $\mathcal{D}$-classes of these monoids and those of previous sections is that here the number of $\mathcal{L}$-classes in D_r is greater than the number of $\mathcal{R}$-classes, while in the previous sections it was the other way around (consider Figures 4.1 and 5.1 for example).

Now we determine where the idempotents are by identifying the group $\mathcal{H}$-classes.

Lemma 5.14. *The $\mathcal{H}$-class $H_{(\alpha,A)}$ in $\text{End}(\text{Aff}(V))$ is a group if and only if H_A is a group in $\text{End}(V)$.*

Proof. The $\mathcal{H}$-class $H_{(\alpha,A)}$ is a group if and only if (α, A) is $\mathcal{H}$-related to $(\alpha, A)^2$ in $\text{End}(\text{Aff}(V))$. This is, by Lemma 5.13, equivalent to saying that A is $\mathcal{H}$ related to A^2 in $\text{End}(V)$ and $\alpha A = (\alpha A+\alpha) - \alpha \in \text{im}\, A$. The first condition is equivalent to saying that H_A is a group in $\text{End}(V)$ and the second is always true. □

Let $N = \{(0, A) \in \text{End}(\text{Aff}(V)) : A \in \text{End}(V)\}$ where 0 denotes the zero vector. Observe that N is a subsemigroup of $\text{End}(\text{Aff}(V))$ and is isomorphic to $\text{End}(V)$. In the same way that T_n dictated the structure of $G \wr T_n$, and $\text{End}(U)$ related to that of $\text{End}(V[W])$, the subsemigroup N will determine the structure of $\text{End}(\text{Aff}(V))$.

It follows from Lemma 5.13 that the $\mathcal{L}$-classes of D_r are in one-one correspondence with the cosets of the subspaces of V of dimension r. The vector space V acts on the $\mathcal{L}$-classes of $\text{End}(\text{Aff}(V))$ in the following way. Let L be an $\mathcal{L}$-class of $\text{End}(\text{Aff}(V))$ in D_r and let $v \in V$. Then define

$$L \cdot v = L(v, I) = \{(w, A)(v, I) : (w, A) \in L\}$$

where I is the identity matrix. Since $(w, B)(v, I) = (w + v, B)$ the set $L \cdot v$ is indeed an $\mathcal{L}$-class in D_r. We will show that the $\mathcal{L}$-classes in $N \cap D_r$ form a transversal of the orbits of the $\mathcal{L}$-classes, that each of the orbits has the same size, and that all the $\mathcal{L}$-classes in a given orbit look the same, in terms of the position of idempotents (i.e. that the map $f : L \to L \cdot v$ defined by $x \mapsto x(v, I)$ sends group $\mathcal{H}$-classes to group $\mathcal{H}$-classes).

Lemma 5.15. *The $\mathcal{L}$-classes in $N \cap D_r$ form a transversal of the orbits of the $\mathcal{L}$-classes of D_r.*

Proof. Let $(\alpha, A) \in D_r$. Then $(\alpha, A) \cdot (-\alpha) = (0, A) \in N$ which means that every orbit contains at least one $\mathcal{R}$-class of $N \cap D_r$.

Suppose that $L_{(0,A)}$ and $L_{(0,B)}$ belong to the same orbit. Then we have

$$L_{(0,A)} = L_{(0,B)} \cdot v = L_{(v,B)}.$$

We conclude that $(0, A)\mathcal{L}(v, B)$ which means, by Lemma 5.13, that $A\mathcal{L}B$ in $\text{End}(V)$ and, as a consequence of Lemma 5.13 and of the fact that $0 \in \text{im}\, A$, it follows that $L_{(0,A)} = L_{(0,B)}$. □

Lemma 5.16. *Let $(v, A), (w, I) \in D_r$. Then $H_{(v,A)}$ is a group $\mathcal{H}$-class if and only if $H_{(v,A)(w,I)}$ is a group $\mathcal{H}$-class.*

Proof. By Lemma 5.14, the $\mathcal{H}$-class $H_{(v,A)}$ is a group if and only if $H_{(v,A)(w,I)} = H_{(v+w,A)}$ is a group. □

Lemma 5.17. *If $(\alpha, A) \in D_r$ then $|\mathcal{O}_{(\alpha,A)}| = q^{n-r}$.*

Proof. From Lemma 5.13 it follows that (α, A) and (β, B) are $\mathcal{L}$-related if and only if $\operatorname{im} A = \operatorname{im} B$ and α and β belong to the same coset of $\operatorname{im} A$ in V. Also, for any orbit $\mathcal{O}$ and any $v \in V$ there exists some element in $\mathcal{O}$ with the form $L_{(v,A)}$. This is because, by Lemma 5.15, every orbit contains some $\mathcal{L}$-class of the form $L_{(0,B)}$ and $(0, B)(v, I) = (v, B)$ which implies that $L_{(v,B)} \in \mathcal{O}$. From this we conclude that each orbit has size

$$[V : \operatorname{im} A] = \frac{|V|}{|\operatorname{im} A|} = \frac{q^n}{q^r} = q^{n-r}.$$

□

Tying together the lemmas above we obtain:

Proposition 5.18. *Let V be a finite vector space of dimension n over the finite field F, with $|F| = q$. Let*

$$I(n, q, r) = \{(v, A) \in \operatorname{End}(\operatorname{Aff}(V)) : \dim(\operatorname{im} A) \leq r\}$$

and

$$I'(n, q, r) = \{A \in \operatorname{End}(V) : \dim(\operatorname{im} A) \leq r\}.$$

Also let $V_{n,r} = I(n, q, r)/I(n, q, r-1)$, $Q_{n,r} = I'(n, q, r)/I'(n, q, r-1)$ and

$$D_r = \{(v, A) \in \operatorname{End}(\operatorname{Aff}(V)) : \dim(\operatorname{im} A) = r\}$$

for $0 \leq r \leq n$. Then we have the following.

(i) *The number of $\mathcal{L}$-classes in D_r is $q^{n-r}\begin{bmatrix} n \\ r \end{bmatrix}_q$.*

(ii) *The number of $\mathcal{R}$-classes in D_r is $\begin{bmatrix} n \\ r \end{bmatrix}_q$.*

(iii) *Let T_1 and T_2 be the 0-rectangular band homomorphic images of $V_{n,r}$ and $Q_{n,r}$ respectively. Then $T_1 \cong {R^0}_{q^{n-r}} \times_0 T_2$ where ${R^0}_k$ denotes the k-element right zero semigroup with a zero adjoined.*

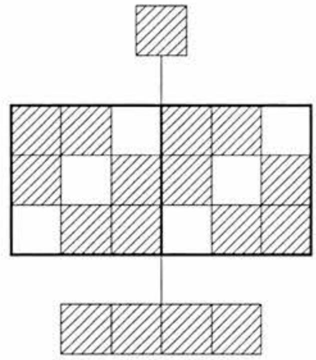


Figure 5.4: Egg-box picture of End(Aff(V)) where $V = \mathbb{Z}_2 \oplus \mathbb{Z}_2$. Here $n = 2$, $q = 2$ and when $r = 1$ we have $q^{n-r} = 2$ and the middle $\mathcal{D}$-class is constructed from 2 copies of the middle $\mathcal{D}$-class of the semigroup End(V) sitting side-by-side.

Proof. (i) Since N is isomorphic to End(V) it follows that the number of $\mathcal{L}$-classes in $N \cap D_r$ is equal to the number of r-dimensional subspaces of V which is given by $\begin{bmatrix} n \\ r \end{bmatrix}_q$. Now the result follows from Lemmas 5.15 and 5.17.

(ii) It follows from Lemma 5.13 that the $\mathcal{R}$-classes are in one-one correspondence with the $(n-r)$-dimensional subspaces of V. This number is given by the expression in the proposition.

(iii) This follows from Lemmas 5.14, 5.15 and 5.17. □

Example 5.19. Let $V = \mathbb{Z}_2 \oplus \mathbb{Z}_2$. Then the semigroup End(Aff(V)) has 3 $\mathcal{D}$-classes, the middle $\mathcal{D}$-class is a horizontal tiling of two copies of the middle $\mathcal{D}$-class of End(V) and the bottom $\mathcal{D}$-class is a horizontal tiling of 4 copies of the bottom $\mathcal{D}$-class of End(V) (see Figure 5.4).

We now calculate the idempotent rank of the ideals of the semigroups End(Aff(V)).

Theorem 5.20. *Let* $S = \text{End}(\text{Aff}(V))$ *with* V *a vector space of dimension* n *over the finite field* F *and with* $|F| = q$. *Let*

$$I(n, q, r) = \{(A, X) \in \text{End}(\text{Aff}(V)) : \dim(\text{im}\, X) \leq r\}.$$

Then we have

$$\text{idrank}(I(n, q, r)) = \text{rank}(I(n, q, r)) = q^{n-r} \begin{bmatrix} n \\ r \end{bmatrix}_q.$$

Proof. By Lemma 4.16 the rank and idempotent rank of $I(n, q, r)$ equals the rank and idempotent rank of $V_{n,r}$. Now the result follows from Lemma 3.6, Proposition 5.18 (Part (iii)) and Theorem 3.39. □

The semigroup $\mathrm{Aff}(V)[+W]$

We saw in Section 5.1 that given a vector space V, and a subspace W of V, we can construct the independence algebra $V[W]$. The endomorphism monoid of this algebra has a structure that is strongly related to that of $\mathrm{End}(U)$. In a similar way, given a subspace W of V we may construct a new non-trivial independence algebra, without constants, by combining $\mathrm{Aff}(V)$ and W. We do not want the new algebra to have any constants, however, so introducing new nullary operations will not work. Start with $\mathrm{Aff}(V)$ and then adjoin a collection of unary operators τ_w for $w \in W$ where

$$x\tau_w = x + w.$$

This new algebra is defined in [14, Example 3.2] where it is denoted by $\mathrm{Aff}(V)[+W]$. The algebra $\mathrm{Aff}(V)[+W]$ is an independence algebra with dimension $n - m + 1$. Its non-empty subalgebras are the affine subspaces of V that contain some coset of W as a subset.

Example 5.21. Recall the example $\mathrm{Aff}(V)$ where $V = \mathbb{Z}_3 \oplus \mathbb{Z}_3$. Let $W = \{(0,0), (1,0), (2,0)\}$, a 1-dimensional subspace of the vector space V. Then $\mathrm{Aff}(V[+W])$ is a 2 dimensional independence algebra. The new unary operations are $\tau_{(0,0)}$, $\tau_{(1,0)}$ and $\tau_{(2,0)}$, and its 1-dimensional subalgebras are the horizontal lines $\{(0,0), (1,0), (2,0)\}$, $\{(0,1), (1,1), (2,1)\}$ and $\{(0,2), (1,2), (2,2)\}$.

The endomorphism monoid of $\mathrm{Aff}(V)[+W]$ is a submonoid of $\mathrm{End}(\mathrm{Aff}(V))$ and consists of all the endomorphisms that preserve the new unary operations $(\tau_w)_{w\in W}$. In other words $(\alpha, A) \in \mathrm{End}(\mathrm{Aff}(V)[+W])$ if and only if for all $w \in W$:

$$\begin{array}{rrcl} & (\tau_w(x))(\alpha, A) & = & \tau_w(x(\alpha, A)) \\ \Leftrightarrow & (x + w)(\alpha, A) & = & \tau_w(xA + \alpha) \\ \Leftrightarrow & (x + w)A + \alpha & = & xA + \alpha + w \\ \Leftrightarrow & xA + wA + \alpha & = & xA + w + \alpha \\ \Leftrightarrow & wA & = & w. \end{array}$$

It follows that

$$\mathrm{End}(\mathrm{Aff}(V)[+W]) = \{(v, A) \in \mathrm{End}(\mathrm{Aff}(V)) : wA = w,\ w \in W\}.$$

Using the representation of elements from $\mathrm{End}(V[W])$ given in Section 5.2 we can identify the elements of $\mathrm{End}(V[+W])$ with the set of pairs:

$$\{(\alpha, (A, X)) : \alpha \in F^n,\ A \in \mathrm{Mat}_{(n-m)\times m}(F),\ X \in \mathrm{GLS}(n-m, F)\}$$

with multiplication

$$(\alpha, (A, X))(\beta, (B, Y)) = (\alpha(B, Y) + \beta,\ (A, X)(B, Y)).$$

As in Section 5.2, let U denote the vector space F^{n-m}. Using identical arguments as in the previous section one may prove:

Lemma 5.22. *Let* $(\alpha, (A, X)), (\beta, (B, Y)) \in \mathrm{End}(\mathrm{Aff}(V[+W]))$. *Then*

(i) $(\alpha, (A, X))\mathcal{L}(\beta, (B, Y))$ *in* $\mathrm{End}(\mathrm{Aff}(V[+W]))$ *if and only if* $X\mathcal{L}Y$ *in* $\mathrm{End}(U)$ *and* $\alpha - \beta \in \mathrm{im}\,(A, X) = \mathrm{im}\,(B, Y)$.

(ii) $(\alpha, (A, X))\mathcal{R}(\beta, (B, Y))$ *in* $\mathrm{End}(\mathrm{Aff}(V[+W]))$ *if and only if* $X\mathcal{R}Y$ *in* $\mathrm{End}(U)$ *and there exists* $Z \in \mathrm{Mat}_{(n-m)\times m}(F)$ *such that* $YZ = A - B$.

(iii) $(\alpha, A)\mathcal{D}(\beta, B)$ *in* $\mathrm{End}(\mathrm{Aff}(V[+W]))$ *if and only if* $\dim(\mathrm{im}\, X) = \dim(\mathrm{im}\, Y)$. □

It follows that the semigroup $\mathrm{End}(V[+W])$ has $n - m + 1$ $\mathcal{D}$-classes given by

$$D_r = \{(\alpha, (A, X)) : \dim(\mathrm{im}\, X) = r\}$$

for $0 \leq r \leq n - m$. Identifying the group $\mathcal{H}$-classes we have:

Lemma 5.23. *The* $\mathcal{H}$*-class* $H_{(\alpha,(A,X))}$ *in* $\mathrm{End}(\mathrm{Aff}(V[+W]))$ *is a group if and only if* $H_{(A,X)}$ *is a group in* $\mathrm{End}(V[W])$ *which is true if and only if* H_X *is a group in* $\mathrm{End}(U)$. □

Define

$$N = \{(0, (A, X)) : (A, X) \in \mathrm{End}(V[W])\} \leq \mathrm{End}(\mathrm{Aff}(V[+W]))$$

which is isomorphic to $\mathrm{End}(V[W])$ and dictates the distribution of idempotents in D_r in the same way as $\mathrm{End}(V)$ did for $\mathrm{Aff}(V)$ in Lemma 5.15. In the same way as for $\mathrm{Aff}(V)$ the vector space V acts on the $\mathcal{L}$-classes of the semigroup $\mathrm{End}(\mathrm{Aff}(V(+W)))$ partitioning them into orbits with

$$L \cdot v = L(v, (0, I)) = \{(w, (A, X))(v, (0, I)) : (w, (A, X)) \in L\}.$$

Using identical arguments to the proofs of Lemmas 5.15 and 5.16 we have the following two results.

Lemma 5.24. *The $\mathcal{L}$-classes in $N \cap D_r$ form a transversal of the orbits of the $\mathcal{L}$-classes of D_r.* □

Lemma 5.25. *Let $(\alpha, (A, X)), (\beta, (0, I)) \in D_r$. Then $H_{(\alpha,(A,X))}$ is a group $\mathcal{H}$-class if and only if $H_{(\alpha,(A,X))(\beta,(0,I))}$ is a group $\mathcal{H}$-class.* □

Computing the size of the orbits gives

Lemma 5.26. *If $(\alpha, (A, X)) \in D_r$ in $\mathrm{End}(\mathrm{Aff}(V[+W]))$ then $|\mathcal{O}_{(\alpha,A)}| = q^{n-m-r}$.*

Proof. From Lemma 5.22 it follows that $(\alpha, (A, X))$ and $(\beta, (B, Y))$ are $\mathcal{L}$-related if $X\mathcal{L}Y$ in $\mathrm{End}(U)$ and α and β belong to the same coset of $\mathrm{im}\,(A, X)$ in the vector space V. Since $\dim(\mathrm{im}\, X) = r$ it follows that $\dim(\mathrm{im}\,(A, X)) = m + r$ and

$$[V : \mathrm{im}\,(A, X)] = \frac{|V|}{|\,\mathrm{im}\,(A, X)|} = \frac{q^n}{q^{m+r}} = q^{n-m-r}.$$

□

It follows that D_r has $q^{n-m-r} \begin{bmatrix} n-m \\ r \end{bmatrix}_q$ $\mathcal{L}$-classes. Moreover, by Proposition 5.8 it follows that D_r has $q^{m(n-m-r)} \begin{bmatrix} n-m \\ r \end{bmatrix}_q$ $\mathcal{R}$-classes. Summarising these results we have the following.

Proposition 5.27. *Let $S = \mathrm{End}(\mathrm{Aff}(V[+W]))$ where V is a finite vector space of dimension n over the finite field F, with $|F| = q$, and W is an m-dimensional subspace of V. Let*

$$I(n, q, r) = \{(v, (A, X)) \in \mathrm{End}(\mathrm{Aff}(V[+W])) : \dim(\mathrm{im}\, X) \leq r\}$$

and

$$I'(n, q, r) = \{A \in \mathrm{End}(U) : \dim(\mathrm{im}\, A) \leq r\}.$$

Also let $V_{n,r} = I(n, q, r)/I(n, q, r-1)$, $Q_{n,r} = I'(n, q, r)/I'(n, q, r-1)$ and

$$D_r = \{(v, (A, X)) \in \mathrm{End}(\mathrm{Aff}(V[+W])) : \dim(\mathrm{im}\, X) = r\}$$

for $0 \leq r \leq n - m$. Then we have the following.

(i) The number of $\mathcal{L}$-classes in D_r is $q^{n-m-r} \begin{bmatrix} n-m \\ r \end{bmatrix}_q$.

(ii) The number of $\mathcal{R}$-classes in D_r is $q^{m(n-m-r)} \begin{bmatrix} n-m \\ r \end{bmatrix}_q$.

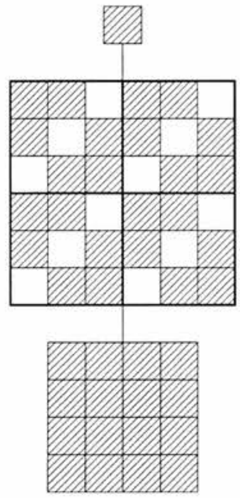


Figure 5.5: Egg-box picture of End(Aff($V[+W]$)) where $V = \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$ and $W = \mathbb{Z}_2 \oplus \{0\} \oplus \{0\}$. In this example $n = 3$, $m = 1$ and $q = 2$ so when $r = 1$ we have $q^{n-m-r} = q^{m(n-m-r)} = 2$. This relates to the fact that the middle $\mathcal{D}$-class in the above diagram is a 2×2 tiling of the middle $\mathcal{D}$-class of End($\mathbb{Z}_2 \oplus \mathbb{Z}_2$).

(iii) Let T_1 and T_2 be the 0-rectangular band homomorphic images of $V_{n,r}$ and $Q_{n,r}$ respectively. Then $T_1 \cong RB^0_{q^{n-m-r},\ q^{m(n-m-r)}} \times_0 T_2$ where $RB^0_{a,b}$ denotes the $a \times b$ rectangular band with a zero adjoined. □

Example 5.28. Let $V = \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$ and $W = \mathbb{Z}_2 \oplus \{0\} \oplus \{0\}$. Then

$$\mathrm{End}(\mathrm{Aff}(V[+W])) \cong \left\{ \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ a & b & c & 0 \\ d & e & f & 0 \\ \hline & \alpha & & 1 \end{array} \right) : a,b,c,d,e,f \in F,\ \alpha \in F^3 \right\}.$$

The algebra Aff($V[+W]$) has dimension 3 and has 3 $\mathcal{D}$-classes. Its egg-box diagram is illustrated in Figure 5.5.

The question of idempotent rank may now be answered for the ideals of the semigroup End(Aff($V[+W]$)).

Theorem 5.29. *Let S = End(Aff($V[+W]$)) where V is a finite vector space of dimension n over the finite field F, with $|F| = q$, and W an m-dimensional subspace of V. Let*

$$I(n,q,r) = \{(v,(A,X)) \in \mathrm{End}(\mathrm{Aff}(V[+W])) : \dim(\mathrm{im}\, X) \leq r\}$$

for $0 \leq r \leq n-m$. *Then we have*

$$\operatorname{idrank}(I(n,q,r)) = \operatorname{rank}(I(n,q,r)) = \begin{cases} q^{m(n-m-r)} \begin{bmatrix} n-m \\ r \end{bmatrix}_q & \textit{if } m \neq 0 \\ q^{(n-m-r)} \begin{bmatrix} n-m \\ r \end{bmatrix}_q & \textit{if } m = 0. \end{cases}$$

Proof. It follows from Proposition 5.27(iii) and Theorem 3.6 that $V_{n,r}$ has an extremal idempotent generating set. The result then follows from Lemma 4.16. □

Note that when $m = 0$ we have $\operatorname{End}(\operatorname{Aff}(V)) \cong \operatorname{End}(\operatorname{Aff}(V[+W]))$ and the above result agrees with Theorem 5.20.

Affine nearfield algebras

The last remaining examples of non-trivial independence algebras are the so called *affine nearfield algebras* (see [14, Example 3.3]). They are constructed in the following way. Let G be a sharply 2-transitive permutation group on a set X. Here 2-transitive means that for any two pairs $(x,y),(z,t) \in X \times X$ there is at least one $g \in G$ such that $g \cdot (x,y) = (g \cdot x, g \cdot y) = (z,t)$. Sharply 2-transitive means that there is precisely one g in G that sends (x,y) to (z,t). Let $\{O_i : i \in I\}$ be the orbits of G on ordered triples of distinct elements of X. Define binary operations for each $i \in I$ by

$$\mu_i(x,x) = x, \ \mu_i(x,y) = z, \text{ if } (x,y,z) \in O_i.$$

The set X along with these binary operations is a 2-dimensional independence algebra whose proper endomorphisms are the constant maps. This means that the only proper ideal is a right zero semigroup and hence has an extremal idempotent generating set.

Proof of Theorem 4.24

Let $\mathcal{A}$ be a finite independence algebra and let $\operatorname{End}(\mathcal{A})$ be the endomorphism monoid of $\mathcal{A}$. By Cameron and Szabó's classification of finite independence algebras the monoid $\operatorname{End}(\mathcal{A})$ must be isomorphic to one of the examples given in either this, or the previous, chapter. Therefore combining the results of this section with Proposition 4.25 and Theorems 4.34, 4.44, 5.11, and 5.20 this completes the proof of Theorem 4.24.

5.4 Proving the result directly from the definition

Before moving on it will be worth our while stopping to reflect on the method of proof employed in order to prove Theorem 4.24. Cameron and Szabó's classification is a highly non-trivial result. In particular, it relies on the classification of finite simple groups. It is not very satisfactory that Theorem 4.24 is proven for all independence algebras basically by checking all of them one at a time. This approach goes against the original philosophy for considering independence algebras. The "nice" thing about independence algebras is the common framework that they provide, allowing results for the general linear semigroup and the full transformation semigroup to be unified with single proofs. What we have resorted to here is, in some sense, exactly the opposite. We began by proving the result for T_n and $\mathrm{End}(V)$ and then built the proof for $\mathrm{End}(\mathcal{A})$ using these two results, together with the classification.

What problems does one come across when trying to prove Theorem 4.24 directly from the definition of independence algebra? In light of the results of Chapter 3 is would seem reasonable to try and prove that every principal factor of $\mathrm{End}(\mathcal{A})$ has a subsquare that is connected and has a uniform distribution of idempotents with respect to some perfect matching. It may be seen that this is indeed the case by glancing through the structural results of the previous two chapters. One problem in proving such a result directly from the definition is that sometimes the number of $\mathcal{R}$-classes dominates the number of $\mathcal{L}$-classes in D_r, and other times it is the other way around. As a result it is unclear whether pinning down a subsquare involves finding an injection from the set of subspaces of dimension r into the set of kernels with weight r, or vice versa. An alternative approach to proving the result may be to prove the result by induction on the dimension of the algebra. This is, for example, how Fountain and Lewin proved that the ideals of $\mathrm{End}(\mathcal{A})$ are idempotent generated.

Open Problem 7. Prove Theorem 4.24 directly from the definition of independence algebra (i.e. without using the classification of Cameron and Szabó).

Chapter 6

Large completely simple subsemigroups and graph colouring

6.1 Introduction

It is well known, by Cayley's theorem, that every finite semigroup S may be embedded in a finite full transformation semigroup T_n. An obvious question to ask is what is the smallest n such that S can be embedded in T_n? In other words, determine

$$\mu(S) = \min\{n \in \mathbb{N} : S \hookrightarrow T_n\}.$$

The corresponding problem for groups of finding minimal faithful permutation representations has been well studied (see [63] and [47] for example). Related to the problem of finding $\mu(S)$ is the question of finding maximal order subsemigroups of T_n. Let $\mathcal{C}$ be a class of finite semigroups and let $S \in \mathcal{C}$. If

$$\nu(n) = \max\{|T| : T \in \mathcal{C},\ T \hookrightarrow T_n\}$$

and if $\nu(m) \leq |S| \leq \nu(m+1)$ then $\mu(S) \geq m$. Of course, in general this lower bound will not be attained.

The question of finding maximal, with respect to inclusion, subsemigroups has been extensively studied. Liebeck, Praeger and Saxl considered maximal subgroups of A_n and S_n in [70]. Yang considered the maximal subsemigroups of various transformation semigroups in [99], [98] and [100]. The results of Yang are, in some sense, special cases of a general description of maximal subsemigroups of finite semigroups given by Graham, Graham and Rhodes in [40]. In [69] Levi and Wood describe a class of maximal subsemigroups of the infinite Baer–Levi semigroup $BL(p,q)$ and conjecture that every maximal subsemigroup of $BL(p,q)$ is one of that type. In [78] Reilly describes a large class of maximal inverse subsemigroups of T_X (where X is infinite).

In terms of determining the maximal order of subsemigroups of S_n and T_n, less is known. Of course we need only look amongst the maximal subsemigroups in order to find those with maximal order, but there still might be many of these to consider. In [13] Burns and Goldsmith determined the maximal orders of the abelian subgroups of the symmetric group S_n. Also, in [100] Yang gives formulae for the cardinalities of the maximal subsemigroups of the semigroup of all singular transformations Sing_n.

The class of completely simple semigroups will be considered in this chapter. We will determine the maximal order completely simple subsemigroups in each $\mathcal{D}$-class of T_n. In particular when $r \geq 3$ we show that the maximal order completely simple subsemigroups in D_r are all left groups (direct products of a left zero semigroups with groups). The case when $r = 2$ is dealt with separately.

The structural information of Proposition 2.14 will be used throughout. Given a subset X of T_n we use $\operatorname{Kers} S$ and $\operatorname{Ims} X$ to denote the set of all kernels and all images, respectively, of elements of X. In other words

$$\operatorname{Kers} X = \{\ker \alpha : \alpha \in X\}, \quad \operatorname{Ims} X = \{\operatorname{im} \alpha : \alpha \in X\}.$$

It follows from Proposition 2.14 that a subsemigroup S of T_n in D_r is completely simple if and only if every image in $\operatorname{Ims} S$ is a transversal of every kernel in $\operatorname{Kers} S$. Given $\alpha \in T_n$ we call $|\operatorname{im} \alpha|$ the *rank* of the transformation α.

We will also need the following well known inequality. Given a set of real numbers $\{a_1, \ldots, a_m\}$ the *arithmetic mean* of this set of numbers is $(a_1 + \ldots + a_m)/m$ and the *geometric mean* is given by $(a_1 a_2 \ldots a_m)^{1/m}$. These two quantities are related by the following inequality:

$$\frac{(a_1 + \ldots + a_m)}{m} \geq (a_1 a_2 \ldots a_m)^{1/m}.$$

This is known as the *arithmetic–geometric means inequality.*

In §6.2 we consider the largest order of left and right zero subsemigroups of T_n. Completely simple semigroups are the subject of §6.3 and it is in this section that the main results of the chapter are presented. Later in §6.3 a full description of all largest order completely simple subsemigroups of T_n is given.

6.2 Left and right zero semigroups

Before we consider the question of the largest order completely simple subsemigroup in T_n we first consider the related problem of finding the largest order left and right zero subsemigroups of T_n.

We start by describing the largest left and right zero semigroups in the $\mathcal{D}$-class D_r. Then to find the largest overall just involves maximizing this number with r in the range $1 \leq r \leq n$.

Proposition 6.1. *Let U be a subsemigroup of T_n where every $\alpha \in U$ satisfies $|\operatorname{im} \alpha| = r$.*

(i) If U is a left zero semigroup then $|U| \leq r^{n-r}$.

(ii) If U is a right zero semigroup then $|U| \leq \lceil n/r \rceil^t \lfloor n/r \rfloor^{r-t}$ where $n \equiv t \mod r$.

Moreover, there exist left and right zero semigroups that attain these bounds.

Proof. (i) Let $A \subseteq X_n$ with $|X| = r$. We will count the number of idempotents in D_r with image X. These are precisely the elements

$$E = \{\epsilon \in T_n : i\epsilon = i \text{ for } i \in A\}.$$

It is clear that $|E| = r^{n-r}$ since we may map each $i \in X_n \setminus A$ anywhere in A.

(ii) It suffices to find the largest set of idempotents with a given kernel of weight r. The difference with the previous case is that, in general, different choices of kernel will give rise to right zero semigroups with different sizes. Our task is to chose a kernel that corresponds to a right zero semigroup of largest possible size.

If $K = \bigcup_{i \in I} K_i$ is a kernel of weight r (i.e. $|I| = r$) then the total number of idempotents with kernel K is equal to $\prod_{i \in I} |K_i|$ (i.e. the number of distinct transversals of the kernel classes $\bigcup_{i \in I} K_i$). It follows that we must determine the number

$$M(n, r) = \max\{a_1 \dots a_r : n = a_1 + \dots + a_r\}$$

over all possible partitions of n into r numbers $a_1, \dots, a_r$.

Claim. $M(n, r) = \lceil n/r \rceil^t \lfloor n/r \rfloor^{r-t}$ *where* $n \equiv t \mod r$.

Proof. If r divides n then from the arithmetic–geometric means inequality

$$(a_1 a_2 \dots a_r)^{1/r} \leq (a_1 + \dots + a_r)/r = n/r$$

which implies that $M(n, r) \leq (n/r)^r$. Also, $M(n, r) \geq (n/r)^r$ just by setting $a_1 = a_2 = \dots = a_r = n/r$.

Now suppose that r does not divide n. Write $n = ar + t$ where $0 < t < r$.

($\geq$) We have $\lfloor n/r \rfloor = a$ and $\lceil n/r \rceil = a + 1$. Now let $a_1 = \dots = a_t = \lceil n/r \rceil$ and $a_{t+1} = \dots = a_r = \lfloor n/r \rfloor$ so that:

$$a_1 + \dots + a_r = t\lceil n/r \rceil + (r - t)\lfloor n/r \rfloor = t(a + 1) + (r - t)a = ar + t = n$$

and

$$a_1 \dots a_r = \lceil n/r \rceil^t \lfloor n/r \rfloor^{r-t}$$

as required.

($\leq$) Let $b_1, \dots, b_r \in \mathbb{N}$ such that $b_1 + \dots + b_r = n$ and $b_1 \dots b_r = M(n, r)$. Since r does not divide n it follows that $r \neq 1$. If $b_1 > \lceil n/r \rceil$ then there is some $b_i \leq \lfloor n/r \rfloor$

with $i \neq 1$ since otherwise for all b_i with $i \neq 1$ we would have $b_i > \lfloor n/r \rfloor \geq \lceil n/r \rceil$. Then this would imply:

$$n = b_1 + \ldots + b_r > \lceil n/r \rceil + (r-1)(\lceil n/r \rceil) = (a+1)r$$

which is a contradiction.

Without loss of generality suppose that $b_2 \leq \lfloor n/r \rfloor$. It follows that

$$b_1 > \lceil n/r \rceil > \lfloor n/r \rfloor \geq b_2$$

implying that $b_1 \geq b_2 + 2$ and so

$$(b_1 - 1)(b_2 + 1) = b_1 b_2 - b_2 + b_1 - 1 \geq b_1 b_2 - b_2 + (b_2 + 2) - 1 = b_1 b_2 + 1 > b_1 b_2.$$

Then we have

$$(b_1 - 1)(b_2 + 1) b_3 \ldots b_r > b_1 \ldots b_r = M(n, r)$$

which is a contradiction. If $b_1 < \lfloor n/r \rfloor$ then we obtain a contradiction using a dual argument. We conclude that $b_i \in \{\lfloor n/r \rfloor, \lceil n/r \rceil\}$ for all i.

Now the result follows from the fact that n may be written in a unique way as a sum, with r terms, of numbers from the set $\{\lfloor n/r \rfloor, \lceil n/r \rceil\}$. □

The result now follows from the claim above.

□

Using the proposition above we may determine the largest left and right zero semigroups in the whole of T_n.

Theorem 6.2. *The largest size of a left zero semigroup in T_n is given by $\max\{u^{n-u} : u \in \{\lfloor x \rfloor, \lceil x \rceil\}\}$ where $x \in \mathbb{R}$ is the solution to $x(1 + \ln x) = n$.*

Proof. Differentiating we obtain:

$$\frac{d}{dx}(x^{n-x}) = x^{n-x+1}(n - x(1 + \ln x)).$$

If u is a solution to the equation $x(1 + \ln x) = n$ then x^{n-x} is increasing when $1 \leq x < u$ and decreasing when $x > u$. Therefore the function x^{n-x} has a maximum when $x(1 + \ln x) = n$. □

Theorem 6.3. *Let U be a right zero subsemigroup of T_n of largest size where $n \geq 2$. Then the size of U is given by:*

$$|U| = \begin{cases} 3^{(n/3)} & \textit{if } n \equiv 0 \mod 3 \\ 4.3^{(n-4)/3} & \textit{if } n \equiv 1 \mod 3 \\ 2.3^{(n-2)/3} & \textit{if } n \equiv 2 \mod 3. \end{cases}$$

Moreover, U is contained in $D_{n/3}$ if $n \equiv 0 \mod 3$ and in $D_{\lceil n/3 \rceil}$ if $n \equiv 2 \mod 3$. If $n \equiv 1 \mod 3$ then there is a largest order right zero semigroup in both $D_{\lceil n/3 \rceil}$ and $D_{\lfloor n/3 \rfloor}$.

Proof. When $n = 2, 3$ the result is easily verified, so suppose that $n \geq 4$. We call any k-tuple $(a_1, \ldots, a_k)$ such that $n = a_1 + \ldots + a_k$, and $1 \leq k \leq n$, a *partition* of n. We will call a partition $(a_1, \ldots, a_k)$ *maximal* if

$$a_1 a_2 \ldots a_k = \max\{a_1 a_2 \ldots a_l : a_1 + a_2 + \ldots + a_l = n,\ 1 \leq l \leq n\}.$$

We define the *value* of a partition to be the product of the terms in the partition. Let $(a_1, \ldots, a_k)$ be a maximal partition of n.

Claim. *$a_i \neq 1$ for all i.*

Suppose without loss of generality that $a_1 = 1$ and look for a contradiction. If $k = 2$ then the partition $(a_1 + a_2) = (1 + a_2)$ has value $1 + a_2$ which is strictly greater than $a_1 a_2 = a_2$, the value of the partition (a_1, a_2). This contradicts the fact that $(a_1, \ldots, a_k)$ is maximal. If $k > 2$ then the partition $(a_1 + a_2, a_3, \ldots, a_k)$ has value

$$(a_1 + a_2) a_3 \ldots a_k = a_3 a_4 \ldots a_k + a_1 a_2 \ldots a_k > a_1 a_2 \ldots a_k$$

which again contradicts the maximality of the partition $(a_1, \ldots, a_k)$. □

Claim. *$a_i \leq 4$ for all i.*

Suppose without loss of generality that $a_1 \geq 5$ and look for a contradiction. Then

$$2(a_1 - 2) a_2 \ldots a_k = (2a_1 - 4) a_2 \ldots a_k > (2a_1 - a_1) a_2 \ldots a_k = a_1 a_2 \ldots a_k$$

which is a contradiction of the maximality of the original decomposition. □

Claim. *If $a_i = 2$ then $a_j \neq 4$ for all j.*

If $a_1 = 2$ and $a_2 = 4$ then

$$(a_1 + 1)(a_2 - 1)a_3 \dots a_k = 9a_3 \dots a_k > 8a_3 \dots a_k = a_1 a_2 a_3 \dots a_k$$

which is a contradiction of the maximality of the original decomposition. □

Similarly if $a_i = 4$ then $a_j \neq 2$ for all j. Using similar arguments it is clear that the number 2 appears at most twice in $(a_1, \dots, a_k)$. Also, the maximum number of occurrences of the number 4 is 1. We conclude that:

(i) if $n \equiv 0 \mod 3$ then $(a_1, \dots, a_k) = (3, \dots, 3)$;

(ii) if $n \equiv 1 \mod 3$ then either $(a_1, \dots, a_k) = (3, \dots, 3, 4)$ or $(a_1, \dots, a_k) = (3, \dots, 3, 2, 2)$ (or some permutation of these k-tuples);

(iii) if $n \equiv 2 \mod 3$ then $(a_1, \dots, a_k) = (3, \dots, 3, 2)$ (or some permutation of this k-tuple).

□

In general the largest right zero and largest left zero semigroups do not lie in the same $\mathcal{D}$-class of T_n. For example, in T_3 the largest right zero semigroup is unique and consists of the set of constant mappings. On the other hand, the largest left zero semigroups can be found in D_2. For example:

$$\left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 2 \end{pmatrix} \right\}$$

is a left zero semigroup of T_3 with largest possible order.

6.3 Completely simple semigroups

In this section we determine the largest order that a completely simple subsemigroup of T_n can have. The results of the previous section provide us with bounds on the size of such a semigroup. Indeed, if U is a left, or right, zero subsemigroup of T_n then $V = \{\alpha \in T_n : \operatorname{im} \alpha \in \operatorname{Ims} U \ \&\ \ker(\alpha) \in \operatorname{Kers} U\}$ is completely simple. It is just the union of the group $\mathcal{H}$-classes that U intersects. Also, if U contains elements of rank r then $|V| = r!|U|$. Therefore a completely simple subsemigroup T of T_n with elements of rank r, and largest possible size, satisfies

$$|T| \geq \max(r!\, r^{n-r},\ r! \lceil n/r \rceil^t \lfloor n/r \rfloor^{r-t})$$

(which equals $r!\, r^{n-r}$ when $r \geq 2$) where $n \equiv t \mod r$. Also, if $r = 1$ then the largest completely simple semigroup has size n (i.e. is the set of constant maps). On the other hand, the number of $\mathcal{L}$-classes in T is at most $\lceil n/r \rceil^t \lfloor n/r \rfloor^{r-t}$ and the number of $\mathcal{R}$-classes is at most r^{n-r}. Therefore

$$|T| \leq r! \lceil n/r \rceil^t \lfloor n/r \rfloor^{r-t} r^{n-r}.$$

We will now show that it is, in fact, the lower of these two bounds that is best possible.

We will convert the problem into a problem concerning counting the number of colourings of a particular graph.

Definition 6.4. Let $\mathcal{A}$ be a nonempty set of r-element subsets of $\{1, \ldots, n\}$. Let $\Gamma(\mathcal{A})$ be that graph with set of vertices $\{1, \ldots, n\}$ and i adjacent to j if and only if $\{i, j\} \subseteq A$ for some $A \in \mathcal{A}$.

By an r-colouring of the graph $\Gamma(\mathcal{A})$ we mean an assignment of r colours to the vertices of the graph such that no two adjacent vertices are coloured with the same colour. Let $C = \{c_1, \ldots, c_r\}$ denote the set of colours. Then formally we define an r-colouring of $\Gamma(\mathcal{A})$ to be a surjective map θ from the vertices of $\Gamma(\mathcal{A})$ onto C satisfying:

$$\forall (i, j) \in E(\Gamma(\mathcal{A})),\ \theta(i) \neq \theta(j).$$

We will use $C_r(\Gamma)$ to denote the set of all r-colourings of a graph Γ. The connection between the graph $\Gamma(\mathcal{A})$ and our problem is given in the following lemma.

Lemma 6.5. *Let U be a completely simple subsemigroup of T_n contained in D_r. Then $|U| \leq |\operatorname{Ims} U||C_r(\Gamma(\operatorname{Ims} U))|$. Also, there exist completely simple semigroups with this size.*

Proof. For every $i \in \operatorname{Ims} U$ let K_i denote the set of kernels that the image i is a transversal of. Then define $\mathcal{K} = \bigcap_{i \in \operatorname{Ims} U} K_i$: the set of all kernels for which every image of $\operatorname{Ims} U$ is a transversal. Clearly we have $|U| \leq |\operatorname{Ims} U| r! |\mathcal{K}|$. The kernel of an r-colouring κ of $\Gamma(\operatorname{Ims} U)$ is a partition of X_n (i.e. the colour classes of the colouring). By the definition of $\Gamma(\operatorname{Ims} U)$ every $I \in \operatorname{Ims} U$ is a transversal of these colour classes. There are $r!$ distinct colourings corresponding to each kernel. On the other hand, every $\kappa \in \mathcal{K}$ gives rise to $r!$ colourings of $\Gamma(\operatorname{Ims} U)$. Thus $r!|\mathcal{K}| = |C_r(\Gamma(\operatorname{Ims} U))|$ and so

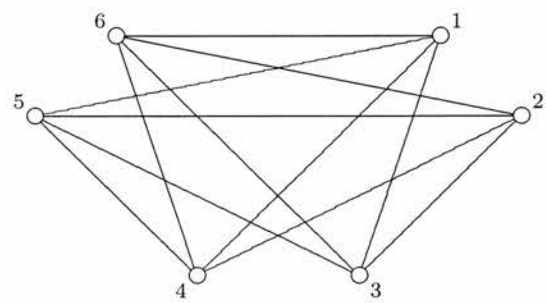
$$|U| \leq |\operatorname{Ims} U| r! |\mathcal{K}| = |\operatorname{Ims} U||C_r(\Gamma(\operatorname{Ims} U))|$$

where equality is achieved by taking all maps with image in $\operatorname{Ims} U$ and kernel in $\mathcal{K}$. □

Example 6.6. Let $n = 6$ and $r = 3$. Also, let

$$\mathcal{A} = \{\{1,3,5\},\{1,3,6\},\{1,4,5\},\{1,4,6\},\{2,3,5\},\{2,3,6\},\{2,4,5\},\{2,4,6\}\}.$$

The 3-sets of $\mathcal{A}$ correspond to the triangles in the complete tripartite graph below.



There are $3! = 6$ ways of colouring the vertices of this graph with three colours. It follows that if U is a completely simple subsemigroup of T_6 with $\operatorname{Ims} U = \mathcal{A}$ then $|U| \leq |\mathcal{A}||C_3(\Gamma(\mathcal{A}))| = 8.6 = 48$.

The next lemma is vital for the proof of the main result of this section.

Lemma 6.7. *Let $\mathcal{A}$ be a set of subsets of $\{1,\ldots,n\}$ each with size r. If $|\mathcal{A}|$ has at least two elements then there exists a strict subset $\mathcal{A}'$ of $\mathcal{A}$ such that $|\mathcal{A}'||C_r(\Gamma(\mathcal{A}'))| \geq |\mathcal{A}||C_r(\Gamma(\mathcal{A}))|$.*

Proof. If $\Gamma(\mathcal{A})$ has no r-colourings then let $\mathcal{A}'$ be any singleton subset of $\mathcal{A}$ and we have $|\mathcal{A}'||C_r(\Gamma(\mathcal{A}'))| \geq 0 = |\mathcal{A}||C_r(\Gamma(\mathcal{A}))|$. Suppose on the other hand that $\Gamma(\mathcal{A})$ has at least one r-colouring. Since $\mathcal{A}$ has at least two elements and is r-colourable there must exist x, y in X_n such that the degrees of x and y are both non-zero and that x and y are not connected by an edge in $\Gamma(\mathcal{A})$. Indeed, if no such pair x, y existed then every vertex with non-zero degree in $\Gamma(\mathcal{A})$ would be connected to every other vertex with non-zero degree in $\Gamma(\mathcal{A})$. Since $\Gamma(\mathcal{A})$ is r-colourable this would mean that $\Gamma(\mathcal{A})$ has at most r vertices with non-zero degree contradicting that fact that $|\mathcal{A}| \geq 2$. Choose such a pair x, y and suppose without loss of generality that $d(x) \leq d(y)$. Now let $\mathcal{A}'$ consist of all the elements of $\mathcal{A}$ that do not contain the number x. Since $d(x) \leq d(y)$ it follows that $|\mathcal{A}'| \geq |\mathcal{A}|/2$ (i.e. we have discarded no more than half of the sets from $\mathcal{A}$). In the graph $\Gamma(\mathcal{A}')$ vertex x has degree zero. Now consider the number of r-colourings of the graph $\Gamma(\mathcal{A}')$. Let Ψ be the set of all colourings of $\Gamma(\mathcal{A})$ where each $\psi \in \Psi$ is a map from X_n to a set of r distinct colours C. Since $\Gamma(\mathcal{A}')$ is a subgraph of $\Gamma(\mathcal{A})$ every r-colouring of $\Gamma(\mathcal{A})$ is an r-colouring of $\Gamma(\mathcal{A}')$. For each colouring $\psi \in \Psi$ we define the following new r-colouring $\theta_\psi : X_n \to C$ by

$$\theta_\psi(m) = \begin{cases} \psi(m) & \text{if } m \neq x \\ c \in C \setminus \psi(x) & \text{otherwise.} \end{cases}$$

Let $\Theta = \{\theta_\psi : \psi \in \Psi\}$. We claim that $\Psi \cup \Theta$ are colourings of $\Gamma(\mathcal{A}')$ and that they are all distinct. Each of the maps Ψ is a colouring of $\Gamma(\mathcal{A}')$ since $\Gamma(\mathcal{A}')$ is a subgraph of $\Gamma(\mathcal{A})$. Let $\theta_\psi \in \Theta$ and let v, w be adjacent vertices in the graph $\Gamma(\mathcal{A}')$. Since x has degree zero in $\Gamma(\mathcal{A}')$ it follows that $v \neq x$ and $w \neq x$. Therefore

$$\theta_\psi(v) = \psi(v) \neq \psi(w) = \theta_\psi(w)$$

and θ_ψ is an r-colouring. The colourings Ψ are all distinct by assumption. To see that the colourings Θ are distinct note that if $\theta_{\psi_1} = \theta_{\psi_2}$ then ψ_1 and ψ_2 agree on every vertex of $\Gamma(\mathcal{A})$ except possibly on x. However, in the graph $\Gamma(\mathcal{A})$ vertex x has non-zero degree and so is contained in a subgraph that is isomorphic to K_r (the complete graph on r vertices). The colour of x, therefore, is uniquely determined once the rest of the vertices have been coloured. We conclude that if $\theta_{\psi_1} = \theta_{\psi_2}$ then $\psi_1 = \psi_2$ and as a consequence that $|\Theta| = |\Psi|$ and the colourings Ψ are all distinct. Also, if $\theta_{\psi_1} \in \Theta$ is different from $\psi_2 \in \Psi$ then either they differ on some vertex in $X_n \setminus \{x\}$ or they agree on all of these vertices in which case (as above) $\psi_1 = \psi_2$ and $\theta_{\psi_1}(x) \notin C \setminus \psi_1(x) = C \setminus \psi_2(x)$ and thus $\theta_{\psi_1}(x) \neq \psi_2(x)$. It follows that $\Psi \cup \Theta$ are colourings of $\Gamma(\mathcal{A}')$ and that

$$|\mathcal{A}'||C_r(\Gamma(\mathcal{A}'))| \geq |\mathcal{A}'|(|\Psi| + |\Theta|) \geq (\frac{|\mathcal{A}|}{2})(2|C_r(\Gamma(\mathcal{A}))|) = |\mathcal{A}||C_r(\Gamma(\mathcal{A}))|$$

as required. □

In fact, in the argument above if $|C| \geq 3$ then $|C \setminus \psi(x)| \geq 2$ so that for every colouring of the old graph we can define two new colourings of the new graph. Then by exactly the same arguments we have:

$$\begin{aligned} |\mathcal{A}'||C_r(\Gamma(\mathcal{A}'))| &\geq |\mathcal{A}'|(|\Psi| + |\Theta|) \geq (\frac{|\mathcal{A}|}{2})(3|C_r(\Gamma(\mathcal{A}))|) \\ &= \frac{3}{2}|\mathcal{A}||C_r(\Gamma(\mathcal{A}))| > |\mathcal{A}||C_r(\Gamma(\mathcal{A}))|. \end{aligned}$$

Interpreted in terms of maximal order completely simple semigroups these results tell us the following.

Proposition 6.8. *Let U be a completely simple subsemigroup of T_n contained in D_r where $r \geq 2$ and $|\operatorname{Ims} U| \geq 2$. Then these exists a completely simple semigroup $V \subseteq D_r$ such that $|\operatorname{Ims} V| \leq |\operatorname{Ims} U|$ and $|V| \geq |U|$. Moreover, if $r > 2$ then V may be chosen so that $|V| > |U|$.* □

It is important to note that the lemma above does not say we can remove $\mathcal{L}$-classes one at a time with the size of the completely simple semigroup staying at least as large at each step.

Example 6.9. In Example 6.6 if we remove a single element from $\mathcal{A}$ to obtain $\mathcal{A}'$, say $\mathcal{A}' = \mathcal{A} \setminus \{1, 3, 5\}$, then the number of distinct 3-colourings of the graph $\Gamma(\mathcal{A}')$ is still only 3! and it follows that

$$|\mathcal{A}'||C_3(\Gamma(\mathcal{A}'))| = 7.6 < 8.6 = |\mathcal{A}||C_3(\Gamma(\mathcal{A}))|.$$

By symmetry, the same is true regardless of which 3-set we remove from $\mathcal{A}$.

We may now prove the main result of this section.

Theorem 6.10. *Let U be a subsemigroup of T_n of mappings all with rank r where $r \geq 2$. If U is completely simple then $|U| \leq r! r^{n-r}$, and this number is attained by the completely simple semigroup $\{\alpha \in T_n : \operatorname{im} \alpha = \operatorname{im} \alpha^2 = \{1, \ldots, r\}\}$.*

Proof. By repeated application of Lemma 6.8 it follows that for some $I \in \operatorname{Ims} U$ there is a completely simple subsemigroup V of D_r such that $\operatorname{Ims} V = \{I\}$ and $|U| \leq |V|$. It follows from Lemma 6.1 that $|U| \leq |V| \leq r! r^{n-r}$. □

As a consequence of this result, and since

$$n! \geq (n-1)(n-1)! \geq (n-2)^2(n-2)! \geq (n-3)^3(n-3)! \geq \ldots \geq 2^{n-2}2!,$$

we have

Theorem 6.11. *If U is a completely simple subsemigroup of $K(n, r)$ with maximal order then $|U| = r!\ r^{n-r}$. In particular, the largest completely simple subsemigroup in T_n is S_n.*

In general, the maximal order completely subsemigroups of D_2 are not all isomorphic, as the following examples show.

Example 6.12. The elements

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 1 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 3 & 4 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 3 & 4 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 3 & 4 & 4 \end{pmatrix}$$

generate a completely simple semigroup which is a subsemigroup of T_5 with two $\mathcal{L}$-classes, four $\mathcal{R}$-classes and has size $16 = 2!\,2^3$.

Example 6.13. The elements

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 2 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 3 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 3 & 3 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 3 & 2 & 2 \end{pmatrix}$$

generate a completely simple semigroup which is a subsemigroup of T_5 with two $\mathcal{L}$-classes, four $\mathcal{R}$-classes and has size $16 = 2!\,2^3$.

Example 6.14. The elements

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 2 & 1 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 3 & 2 & 2 & 3 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 3 & 4 & 3 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 1 & 4 & 4 & 4 \end{pmatrix}$$

generate a completely simple semigroup which is a subsemigroup of T_5 with four $\mathcal{L}$-classes, four $\mathcal{R}$-classes and has size $32 = 2!\,2^4$.

In the example above we have an example of a largest completely simple semigroup with four $\mathcal{L}$-classes. We will now show that this is in fact the largest number of $\mathcal{L}$-classes that a maximal order completely simple semigroup in D_2 can have (for any value of n).

The maximal (with respect to inclusion) subsemigroups of T_n in D_2 are in one-one correspondence with the set of all, not necessarily connected, bipartite graphs on n vertices. In Figure 6.1 all of the maximal completely simple subsemigroups of D_2 in T_5 are described using this correspondence. In this example there are four types of maximal order completely simple subsemigroup, each with 16 elements (i.e. the last 4 in the table). We will show that this example is representative of the general situation since every maximal order subsemigroup in D_2, for any n, corresponds to one of the types given in Figure 6.2.

Theorem 6.15. *Let U be a completely simple subsemigroup of $K(n, r)$ with maximal order.*

(i) If $r = 1$ then $U = K(n, r)$ which is an n element right zero semigroup.

(ii) If $r = 2$ then U is of one of the 4 types given in Figure 6.2.

(iii) If $r \geq 3$ then U is the direct product of an r^{n-r}-element left zero semigroup and the symmetric group S_r (i.e. is a left group).

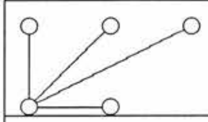
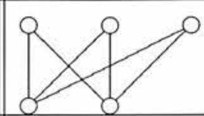
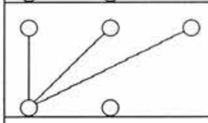
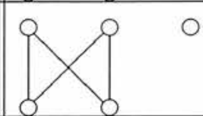
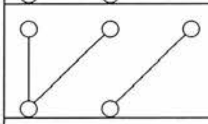
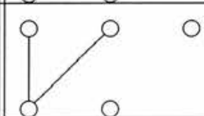
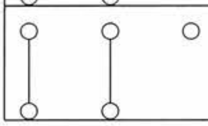
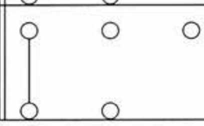
Graph	Dimensions	Size	Graph	Dimensions	Size
	1×4	8		1×6	12
	2×3	12		2×4	16
	2×3	12		4×2	16
	4×2	16		8×1	16

Figure 6.1: The graphs associated with the maximal completely simple semigroups in $D_2 \subseteq T_5$.

Let $\mathcal{A}$ be a set of 2-sets of X_n. The 2-sets of $\mathcal{A}$ are in one-one correspondence with the edges of the graph $\Gamma(\mathcal{A})$. The number of 2-colourings of the graph is simply a function of the number of connected components of the graph $\Gamma(\mathcal{A})$. Indeed, we have $C_2(\Gamma(\mathcal{A})) = 2^k$ where k is the number of connected components in the graph $\Gamma(\mathcal{A})$. We call the connected components of $\Gamma(\mathcal{A})$ with only a single vertex the *trivial* components of the graph.

Lemma 6.16. *Let $U \leq J_2$ be completely simple. If $\Gamma(\operatorname{Ims} U)$ has more than two non-trivial components then there exists a completely simple semigroup $V \subseteq J_2$ such that $|V| > |U|$.*

Proof. The graph $\Gamma(\operatorname{Ims} U)$ is bipartite since it is two colourable. Let $A_i \cup B_i$ where $1 \leq i \leq l$ be the non-trivial connected components of the graph $\Gamma(\operatorname{Ims} U)$ and let $C \subseteq V(\Gamma(\operatorname{Ims} U))$ be the set of all trivial components (i.e. the set of vertices with degree zero). Let e_i denote the number of edges in the component $A_i \cup B_i$ for $1 \leq i \leq l$. Since $\Gamma(\operatorname{Ims} U)$ has at least three non-trivial components we may assume without loss of generality that $e_1 \leq |E|/3$ (since $\sum_{i=1}^{l} e_i = |E|$). Let V be the unique largest completely simple subsemigroup of D_2 with $\operatorname{Ims} V = \operatorname{Ims} U \setminus (A_1 \times B_1)$. We claim that $|V| > |U|$. By construction $|\operatorname{Ims} V| \geq \frac{2}{3}|\operatorname{Ims} U|$ and $|C_2(\Gamma(\operatorname{Ims} V))| \geq 2|C_2(\Gamma(\operatorname{Ims} U))|$ (by exactly the same argument as in the

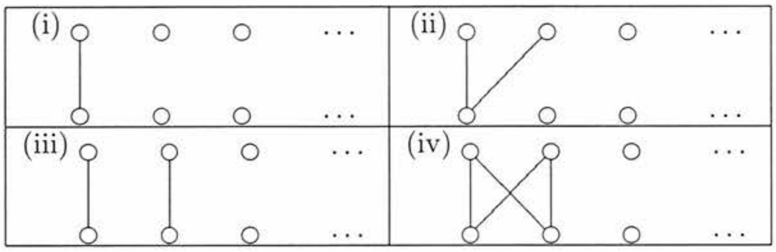


Figure 6.2: The bipartite graphs corresponding to the maximal order completely simple semigroups in D_2.

proof of Lemma 6.7). It follows that

$$\begin{aligned} |V| &= |\operatorname{Ims} V||C_2(\Gamma(\operatorname{Ims} V))| \geq \frac{4}{3}|\operatorname{Ims} U||C_2(\Gamma(\operatorname{Ims} U))| \\ &> |\operatorname{Ims} U||C_2(\Gamma(\operatorname{Ims} U))| = |U| \end{aligned}$$

as required. □

It follows that if we are given a completely simple subsemigroup U of D_2, such that $\Gamma(\operatorname{Ims} U)$ has three or more non-trivial components, then U is definitely not of maximal order in its $\mathcal{D}$-class. As we saw in the examples of Figure 6.1 it is possible to have maximal order examples with two components, but by Theorem 6.11 this is only possible when $r = 2$.

Lemma 6.17. *Let U be a maximal order completely simple subsemigroup of D_2 in T_n where $\Gamma(\operatorname{Ims} U)$ has at least two components. Let E be the set of edges in a largest connected component of $\Gamma(\operatorname{Ims} U)$. Define V to be the unique largest completely simple semigroup with $\operatorname{Ims} V = E$. Then $|V| \geq |U|$ and V is also a maximal order completely simple subsemigroup of D_2.*

Proof. We have

$$|V| = |C_2(\Gamma(\operatorname{Ims} V))||\operatorname{Ims} V| \geq 2|C_2(\Gamma(\operatorname{Ims} U))|\frac{1}{2}|\operatorname{Ims} U| \geq |U|$$

as required. □

Let U be a maximal order completely simple subsemigroup of D_2 such that $\Gamma(\operatorname{Ims} U)$ has only one non-trivial connected component. Let $A \cup B$ be the non-trivial component and let D be the set of vertices with degree zero. Since U has maximal order, and the number of colourings is simply a function of the number of connected components, it follows that every vertex of A is connected to every

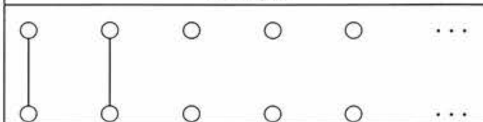
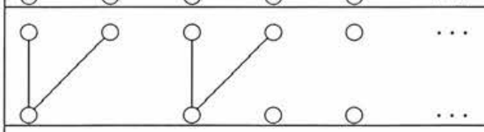
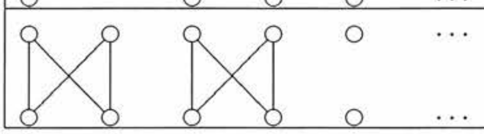
Graph	Size
	$\lvert E\rvert\lvert C\rvert = 2.2^{n-4+2} = 2^{n-1}$
	$\lvert E\rvert\lvert C\rvert = 4.2^{n-6+2} = 2^{n-2}$
	$\lvert E\rvert\lvert C\rvert = 8.2^{n-8+2} = 2^{n-2}$

Figure 6.3: Two component bipartite graphs such that the largest components are maximal and the total number of edges is a power of two.

vertex of B. Thus we have:

$$|U| = |\operatorname{Ims} U||C_2(\Gamma(U))| = ab2^{d+1}$$

where $a = |A|$, $b = |B|$ and $d = |D|$. From Theorem 6.11, since U has maximal order, it follows that $|U| = ab2^{c+1} = 2^{n-1}$. Since $c = n - a - b$ it follows that a, b must satisfy the equation:

$$ab = 2^{a+b-2}$$

where $a, b \in \mathbb{N}$. When $a + b > 8$ we have $ab \leq (a+b)^2 < 2^{(a+b)-2}$. Testing all values in the range $1 \leq a, b \leq 8$ it is easily verified that the only solutions (a, b), with $a \leq b$, are given by

$$(a, b) \in \{(1, 1), (1, 2), (2, 2)\}.$$

These pairs correspond to the graphs (i), (ii) and (iv) in Figure 6.2. This gives all the single non-trivial component solutions to our problem.

Now we consider the two component solutions. Let U be a maximal order completely simple subsemigroup of D_2 such that $\Gamma(\operatorname{Ims} U)$ has exactly two non-trivial connected components. It follows from Lemma 6.17 that either the two non-trivial components of $\Gamma(\operatorname{Ims} U)$ are isomorphic and they both correspond to maximal single component solutions (i.e. those given in Figure 6.2), or they are not isomorphic and the larger of the two components corresponds to a maximal single component solution.

Start with a maximal single component solution (see Figure 6.2) and consider all ways of extending it to a two component solution where the new component has strictly fewer edges than the component we started with. Keep in mind the restriction that the number of edges must be a power of two, for the graph to stand a chance of being maximal, since $|E||C_2| = 2^{n-1}$ and the number of 2-colourings is always a power of two. It is now easily verified, by exhaustion, that only three graphs may arise in this way. They are given in Figure 6.3. From the second column of the table in Figure 6.3 it follows that the only two component maximal solution is given by the graph (iii). Since, by Lemma 6.16, every maximal graph has at most two non-trivial components, this completes the proof of Theorem 6.15

Of course, for other classes of semigroup we can ask the same question. In [44] the largest size of an inverse subsemigroup of T_n is computed. In particular it is shown that:

Theorem 6.18. *The largest possible size of an inverse subsemigroup of T_n is $\sum_{m=0}^{n-1} \binom{n-1}{m}^2 m!$. Moreover, the subsemigroup generated by the transformations:*

$$\begin{pmatrix} 1 & 2 & 3 & \dots & n-1 & n \\ 2 & 1 & 3 & \dots & n-1 & n \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & \dots & n-1 & n \\ 2 & 3 & 4 & \dots & 1 & n \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & \dots & n-1 & n \\ 1 & 2 & 3 & \dots & n & n \end{pmatrix}$$

is an inverse subsemigroup of T_n with this size.

One class closely related to the completely simple semigroups are the completely regular semigroups. These are the semigroups with the property that every element lies in a subgroup of the semigroup. We leave the task of finding large completely regular semigroups as an open problem. Clearly the answer is at least $n!$ but is there are larger example?

Open Problem 8. Determine the maximal order of a completely regular subsemigroup of T_n.

Acknowlegment

The work in this chapter was carried out in collaboration with Dr J. D. Mitchell of the University of St Andrews.

Part II

Infinite semigroup theory

Chapter 7

Generators and relations via boundaries in Cayley graphs

7.1 Introduction

In previous chapters we have been concerned with the problem of finding "nice" generating sets for finite semigroups. We will now move our attention to infinite semigroups. We also change our question from that of the rank of the semigroup to the question of whether or not the semigroup may be generated by a finite set. Such semigroups are called *finitely generated*. Every finite semigroup is, of course, finitely generated. Also, for an infinite semigroup to stand a chance of being finitely generated it must be countably infinite (since the set of all finite words over a finite alphabet is a countable union of finite sets, words of length one, two three, etc..., and hence is countable). It was observed in our rank investigations that the rank function is not well behaved with respect to taking subsemigroups. Analogous to this is the fact that it is possible to embed non-finitely generated semigroups in finitely generated ones (we will see examples of this later in the chapter). In the previous chapters we approached the study of finite semigroups by concentrating mainly on semigroups of transformations, of one form or another. When working with infinite semigroups the theory of semigroup presentations, representing semigroups as factor semigroups of free semigroups, is often a more profitable approach. As a result of this, the theory of semigroup presentations will be central this part of the thesis.

Given a semigroup S and a subsemigroup T of S it is natural to consider which properties S and T have in common. In the case of groups, for example, it is known that a group shares many of its properties with its subgroups of finite index. In particular we have the Reidemeister–Schreier theorem which says that subgroups of finitely presented groups with finite index are finitely presented; see [71, Proposition 4.2]. The general study of subgroups of finitely presented groups continues to receive a lot of attention; see for example [9], [22] and [90]. An important problem in the development of a similar theory for arbitrary monoids has been the search for a suitable notion of index for subsemigroups. One approach is to define the index of T in S to be the cardinality of the set $S \setminus T$. This is normally known as the *Rees* index of T in S. In [61] and [62] Jura discusses the problem of finding all the ideals of a given Rees index in a finitely presented semigroup. In order to obtain this result he proves the Hilbert–Schreier theorem for semigroups i.e. that if S is a finitely generated semigroup and T is a subsemigroup of S with finite Rees index then T is finitely generated. This result was reproved in [82] where, in addition, it was also shown that subsemigroups of finitely presented semigroups with finite Rees index are themselves finitely presented, although the proof given in that paper is incomplete. We discuss how the proof may be fixed in §7.5. An important tool that was used in the proof of this result is the Reidemeister–Schreier rewriting theorem for semigroups introduced in [17].

In [101] and [102] the groups of units of finitely presented monoids are considered. The author considers so called special monoids and proves that, for this class of monoid, from a finite presentation for the monoid we may obtain a finite presentation for the group of units (with the same number of defining relations). In [83] and [84] presentations for arbitrary subgroups of finitely presented monoids are considered. In particular in [83] an example is given of a finitely presented monoid whose group of units is not finitely presented. Presentations of ideals of finitely presented semigroups were considered in [19] and of those of arbitrary subsemigroups were considered in [18].

In [53] automatic semigroups were investigated and it was shown that if T is a finite Rees index subsemigroup of S then S is automatic if and only if T is.

The theory of monoid presentations is closely linked to that of string-rewriting systems. An important problem in this area is to classify all monoids that may be presented by some finite complete string-rewriting system. Monoids that may be defined by such presentations have nice properties: for example they all have solvable word problem. On the other hand, in [88] Squier showed that not every monoid that has solvable word problem is presented by some finite complete string-rewriting system. In a subsequent paper [89] Squier, Otto and Kobayashi introduced the notion of finite derivation type, proving that a monoid has finite derivation type if it can be presented by a finite complete rewriting system. In [93] it was shown that if T is a finite Rees index subsemigroup of S and T has finite derivation type then so does S. The converse of this result is still an open problem although has recently been solved by Malheiro [1] in the special case where T is an ideal with finite Rees index. In the same paper it was also shown that if T has finite Rees index in S then S can be presented by a finite complete rewriting system if T can.

In this chapter we introduce a new notion of index for subsemigroups which is significantly weaker than Rees index but is still strong enough to force T to inherit certain properties from S. The general idea is that rather than forcing the entire complement $S \setminus T$ to be finite we need only restrict the number of points where T and $S \setminus T$ meet each other in the Cayley graphs to be finite.

Let S be a finitely generated semigroup with T a subsemigroup of S. Let A be a finite generating set of S. Let $\Gamma_r(A, S)$ and $\Gamma_l(A, S)$ denote the right and left Cayley graphs of S with respect to A. Thus the vertices of $\Gamma_r(A, S)$ are the elements of S and there is a directed edge from s to t, labelled with $a \in A$, if and only if $sa = t$. The left Cayley graph is defined analogously. We define the *right boundary edges of T in $\Gamma_r(A, S)$* to be those edges whose initial vertex is in $S \setminus T$ and terminal vertex is in T. The left boundary edges are defined in the same way but using the left Cayley graph. We define the *right boundary of T in S with respect to A* to be the set of terminal vertices of the right boundary edges of T in $\Gamma_r(A, S)$ together with the elements of A that belong to T. The *left boundary of T in S with respect to A* is defined to be the set of terminal vertices of the left boundary edges of T in $\Gamma_l(A, S)$ together with the elements of A that belong to T. We define the *(two-sided) boundary* of T in S to be the union of the left and right boundaries. We use $\mathcal{B}_l(A, T)$, $\mathcal{B}_r(A, T)$ and $\mathcal{B}(A, T)$ to denote the left, right and two-sided boundaries, respectively, of T in S with respect to A. Formally these sets are given by

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

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

and

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

where S^1 denotes S with an identity adjoined (even if it already has one), U denotes the complement $S \setminus T$ and U^1 denotes $S^1 \setminus T$. We say that T *has a 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 Proposition 7.3). 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.

The main results of this chapter show that the properties of finite generation and presentability are inherited by subsemigroups with finite boundary.

Theorem 7.1. *If S is a finitely generated semigroup and T is a subsemigroup of S with finite boundary then T is finitely generated.*

Theorem 7.2. *Let S be a finitely generated semigroup and T be a subsemigroup of S. If S is finitely presented and T has a finite boundary in S then T is finitely presented.*

The chapter is structured as follows. We begin by describing the basic prop-

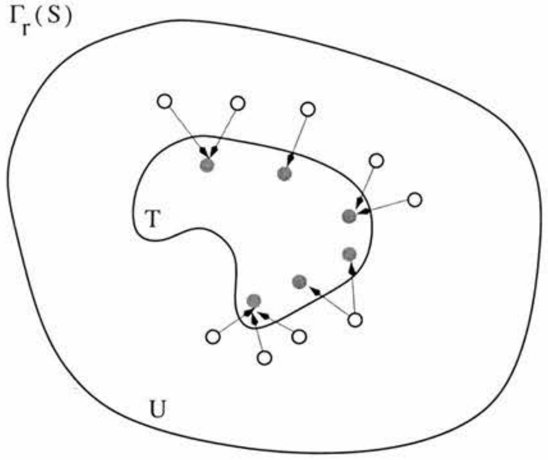


Figure 7.1: Visualising the right boundary of the subsemigroup T of the semigroup S.

erties of boundaries in §7.2. In §7.3 we show the connection between boundaries and generating sets of subsemigroups and in the process prove Theorem 7.1. We consider semigroup presentations in §7.4 and we prove Theorem 7.2. In §7.5 we give some applications of the main results and we show that having a finite boundary is not a transitive property. The question of finite presentability when only one of the boundaries is finite is the subject of §7.6 and in §7.7 the possible sizes of left and right boundaries are explored. The converse of Theorem 7.2 is the subject of §7.8 and in §7.9 boundaries of subsemigroups of free semigroups are investigated.

7.2 Examples and basic properties

Let S be a semigroup generated by a finite set A. Let A^+ denote the set of all non-empty words over the alphabet A, and let A^* denote the set of all words over the alphabet A including the empty word ϵ. There is a natural homomorphism $\theta : A^+ \to S$ mapping each word in A^+ to its corresponding product of generators in S. Since A generates S the map θ is surjective. Associated with the map θ is a congruence η on the free semigroup A^+ given by $(w, v) \in \eta$ if and only if $w\theta = v\theta$. Then the quotient A^+/η is isomorphic to S under the natural map $w/\eta \mapsto w\theta$. Given some $w \in A^+$ we will, where there is no chance of confusion, often omit reference to the function θ or the relation η altogether and talk of w in S rather than $w\theta$ in S or $w/\eta \in A^+/\eta$.

Given a word $w \in A^+$ we will use $|w|$ to denote its length. Given $w, v \in A^+$ we write $w = v$ if they represent the same element of S (i.e. if $w\theta = v\theta$) and write $w \equiv v$ if they are identical as words in A^+. Furthermore, given $w \in A^+$ and $s \in S$ we write $w = s$ meaning that $w\theta = s$ in S. We write S^1 to denote the semigroup S with an identity adjoined (even if S already has an identity) and we extend the definition of θ so that $\theta : A^* \to S^1$ by setting $\epsilon\theta = 1$.

For two words $w, v \in A^+$ we say that w is a *prefix* (respectively suffix) of v if $v \equiv w\beta$ (respectively $v \equiv \beta w$) for some $\beta \in A^*$. We say that w is a *subword* of v if $v \equiv \alpha w \beta$ where $\alpha, \beta \in A^*$. Also, for a subset Y of S^1 we define $\mathcal{L}(A, Y) = \{w \in A^* : w\theta \in Y\}$ and call this set the *language of* Y *in* A^*. Note that from the convention described in the previous paragraph it follows that, for $Y \subseteq S$, we have $\mathcal{L}(A, Y^1) = \mathcal{L}(A, Y) \cup \{\epsilon\}$.

We now show that whether the boundary is finite or not is independent of the choice of generating set.

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.*

Proof. We will prove the first statement only. The second may be proven using a dual argument. For each $b \in B$ let $\pi_A(b) \in A^+$ be some fixed decomposition of b into generators from A so that $b = \pi_A(b)$ in S. Let $m = \max\{|\pi_A(b)| : b \in B\}$ which exists since B is finite. We claim that

$$\mathcal{B}_r(B,T) \subseteq \bigcup_{i=1}^{m} \mathcal{B}_r(A,T)A^i$$

which is a finite set since both A and $\mathcal{B}_r(A,T)$ are finite. To see this first let $t \in \mathcal{B}_r(B,T)$. By the definition of $\mathcal{B}_r(B,T)$ we can write $t = ub = u\pi_A(b) = ua_1 \ldots a_k$ where $u \in U^1$, $b \in B$, $a_i \in A$ for $1 \leq i \leq k$ and $k \leq m$. Let l be the smallest subscript such that $ua_1 \ldots a_l$ belongs to T. It follows that $ua_1 \ldots a_l \in \mathcal{B}_r(A,T)$ and we have

$$t = (ua_1 \ldots a_l)(a_{l+1} \ldots a_k) \in \mathcal{B}_r(A,T)A^{k-l} \subseteq \bigcup_{i=1}^{m} \mathcal{B}_r(A,T)A^i$$

since $k - l \leq m$. $\square$

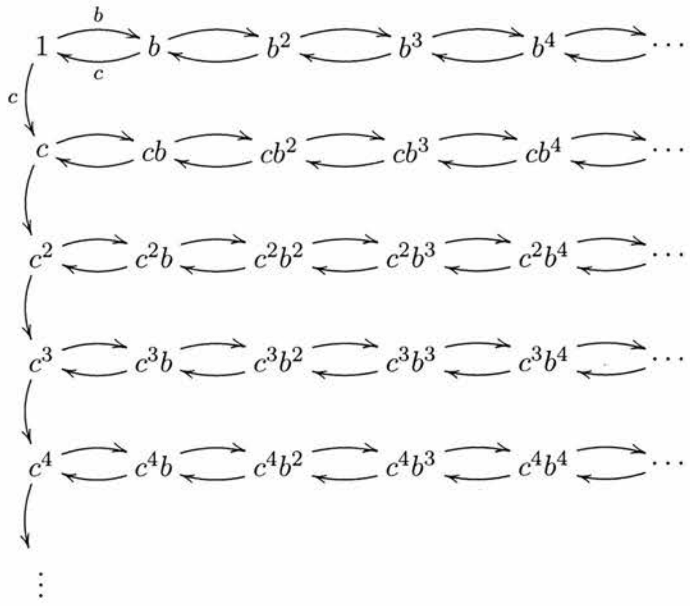


Figure 7.2: The right Cayley graph of the bicyclic monoid.

Example 7.4. Let $B = \mathbb{N}^0 \times \mathbb{N}^0$ subject to the multiplication

$$(m, n)(p, q) = (m - n + \max(n, p), q - p + \max(n, p)).$$

This semigroup is called the *bicyclic monoid* and is generated by $A = \{b, c\}$ where $b = (0, 1)$ and $c = (1, 0)$. The right Cayley graph of B with respect to A is given in Figure 7.2. The subsemigroup $T = \{b^i : i \in \mathbb{N}^0\}$ satisfies $\mathcal{B}_r(A, T) = \{b\}$ and so has a finite right boundary. On the other hand, the subsemigroup $L = \{c^i : i \in \mathbb{N}^0\}$ satisfies $\mathcal{B}_r(A, L) = L$ and hence has an infinite right boundary.

Finite boundaries arise in many natural situations. We describe a few of them in the following proposition. The proof follows straight from the definition of the boundary.

Proposition 7.5. *Let S be a semigroup generated by a finite set A and let T be a subsemigroup of S. Then we have:*

(i) if T is finite then $\mathcal{B}(A, T)$ is finite;

(ii) if $S \setminus T$ is finite then $\mathcal{B}(A, T)$ is finite;

(iii) if $S \setminus T$ is a right (resp. left) ideal in S then $\mathcal{B}_r(A, T)$ (resp. $\mathcal{B}_l(A, T)$) is finite;

(iv) if $S \setminus T$ is an ideal in S then $\mathcal{B}(A,T)$ is finite.

The *depth* of an element $s \in S$ is defined to be the minimal possible length of a product in A^+ that equals s in S, and is denoted by $d(s)$. In other words:

$$d(s) = \min\{|w| : w \in A^+, w = s \text{ in S}\}.$$

For a subset X of S we define the depth of X to be:

$$d(X) = \max\{d(x) : x \in X\}$$

when it exists and say that X has infinite depth otherwise. Also, given a word $w \in A^+$ we define the depth of w by $d(w) = d(w\theta)$.

Fix a transversal $\mathcal{R}$ of the η-classes of A^+ chosen so that every $w \in \mathcal{R}$ is a word of shortest length in its η-class. For every $w \in A^+$ define $\overline{w} = (w/\eta) \cap \mathcal{R}$: the fixed shortest length word in $\mathcal{R}$ that equals w in S.

The next result gives a characterisation of subsemigroups with finite boundary that does not refer to the generating set of S. The properties described in the proposition will be used frequently in later sections.

Proposition 7.6. *Let S be a finitely generated semigroup with T a subsemigroup of S. Then T has a finite boundary in S if and only if the following properties hold:*

(i) for every finite subset X of S the set $U^1X \cap T$ is finite;

(ii) for every finite subset X of S the set $XU^1 \cap T$ is finite;

(iii) the set $U^2 \cap T$ is finite.

Proof. Let A be a finite generating set for S. Suppose that T has an infinite boundary in S. Then either the right or the left boundary must be infinite. Suppose that the right boundary is infinite. Then we can find an infinite subset $\{u_1, u_2, \ldots\}$ of U and a generator $a \in A$ such that $\{u_1a, u_2a, \ldots\}$ is an infinite subset of T. If $a \in U$ then condition (iii) fails. Otherwise $\{a\}$ is a finite subset of T for which condition (i) fails. On the other hand if the left boundary is infinite then, by a dual argument, either condition (iii) or condition (ii) fails.

For the converse suppose that T has a finite boundary in S. We now show that each of the three conditions given in the proposition must hold.

(i) Let X be a finite subset of S. Define $B = X \cup A$ which is a finite generating set for S. Now we have:

$$U^1X \cap T \subseteq U^1B \cap T = \mathcal{B}_r(B,T)$$

where $\mathcal{B}_r(B,T)$ is finite by Proposition 7.3. That condition (ii) holds is proven using a dual argument.

(iii) Let $m = d(\mathcal{B}(A,T))$, the depth of the boundary of T in S, which is well defined since $\mathcal{B}(A,T)$ is finite. Define $Z = \{w \in \mathcal{L}(A,T) : |w| \leq 3m\}$, $Y = (Z\theta) \cap U^2$ and let

$$k = \max_{y \in Y} \min\{|v| : u, v \in \mathcal{L}(A,U), uv = y\}$$

which must exist since Y is finite.

Claim. *For all $u, v \in \mathcal{L}(A,U)$ where $uv \in \mathcal{L}(A,T)$ there exist $u_1, v_1 \in \mathcal{L}(A,U)$ such that $|v_1| \leq k$ and $uv = u_1v_1$ in S.*

Proof. We prove the claim by induction on the length of the word $|uv|$. Let $u, v \in \mathcal{L}(A,U)$ where $uv \in \mathcal{L}(A,T)$. If $|uv| = 2$ then the result holds trivially. Now suppose that the result holds for all pairs $\gamma, \delta \in \mathcal{L}(A,U)$ where $\gamma\delta \in \mathcal{L}(A,T)$ and $|\gamma\delta| < |uv|$. We prove the result for uv by considering the following cases.

Case 1: *u has no prefix in $\mathcal{L}(A,T)$.* In this case since $v \in \mathcal{L}(A,U)$ and $uv \in \mathcal{L}(A,T)$ we can write $uv \equiv u'\beta_1$ where u' is a prefix of u, v is a suffix of β_1 and $\beta_1 \in \mathcal{L}(A, \mathcal{B}_l(A,T))$. We have $uv = u'\overline{\beta_1}$ where, since u has no prefix in $\mathcal{L}(A,T)$, $u' \in \mathcal{L}(A,U^1)$ and $u'\overline{\beta_1} \in \mathcal{L}(A,T)$. We can, therefore, write $u'\overline{\beta_1} = \beta_2\gamma$ where $\beta_2 \in \mathcal{L}(A, \mathcal{B}_r(A,T))$, u' is a prefix of β_2 and γ is a suffix of $\overline{\beta_1}$. Therefore $uv = \beta_2\gamma = \overline{\beta_2}\gamma$ where

$$|\overline{\beta_2}\gamma| = |\overline{\beta_2}| + |\gamma| \leq |\overline{\beta_2}| + |\overline{\beta_1}| \leq 2m \leq 3m.$$

It follows that $d(uv) \leq 3m$ which implies that $(uv)\theta \in Y$ and, by the definition of k, we can write $uv = u_1v_1$ where $|v_1| \leq k$.

Case 2: *u has a prefix in $\mathcal{L}(A,T)$.* In this case, since u has a prefix in $\mathcal{L}(A,T)$, we write $uv \equiv \beta u'v = \overline{\beta}u'v$ where $\beta \in \mathcal{L}(A, \mathcal{B}_r(A,T))$ and, since T is a subsemigroup of S, $u' \in \mathcal{L}(A,U)$. This case now splits into two subcases depending on whether or not $u'v \in \mathcal{L}(A,T)$.

Case 2.1: *$u'v \notin \mathcal{L}(A,T)$.* Since $\overline{\beta}u'v \in \mathcal{L}(A,T)$ we can write $\overline{\beta}u'v \equiv \gamma\beta_2$ where $\beta_2 \in \mathcal{L}(A, \mathcal{B}_l(A,T))$, γ is a prefix of $\overline{\beta}$ and $u'v$ is a suffix of β_2. Now we have

$uv = \overline{\beta}u'v = \gamma\overline{\beta_2}$ which satisfies $|\gamma\overline{\beta_2}| \leq 2m \leq 3m$. It follows that $d(uv) \leq 3m$ which implies that $(uv)\theta \in Y$ and, by the definition of k, we can write $uv = u_1v_1$ where $|v_1| \leq k$.

Case 2.2: $u'v \in \mathcal{L}(A,T)$. In this case $u', v \in \mathcal{L}(A,U)$, $u'v \in \mathcal{L}(A,T)$ and $|u'v| < |uv|$, so we can apply induction writing $u'v = u_2v_2$ where $|v_2| \leq k$. Now we have $uv \equiv \beta u'v = \overline{\beta}u_2v_2$. If $\overline{\beta}u_2 \in \mathcal{L}(A,U)$ then we are done since $uv = (\overline{\beta}u_2)v_2$ where $\overline{\beta}u_2, v_2 \in \mathcal{L}(A,U)$ and $|v_2| \leq k$. On the other hand, if $\overline{\beta}u_2 \in \mathcal{L}(A,T)$ then since $u_2 \in \mathcal{L}(A,U)$ and $\overline{\beta}u_2 \in \mathcal{L}(A,T)$ we can write $\overline{\beta}u_2 \equiv \gamma\beta_1$ where $\beta_1 \in \mathcal{L}(A, \mathcal{B}_l(A,T))$, γ is a prefix of $\overline{\beta}$ and u_2 is a suffix of β_1. Now we have $uv = \beta u'v = \overline{\beta}u_2v_2 = \gamma\overline{\beta_1}v_2$. Since $v_2 \in \mathcal{L}(A,U)$ and $\gamma\overline{\beta_1}v_2 \in \mathcal{L}(A,T)$ we can write $\gamma\overline{\beta_1}v_2 \equiv \delta\beta_2$ where $\beta_2 \in \mathcal{L}(A, \mathcal{B}_l(A,T))$, δ is a prefix of $\gamma\overline{\beta_1}$ and v_2 is a suffix of β_2. Therefore $uv = \delta\overline{\beta_2}$ where

$$|\delta\overline{\beta_2}| \leq |\delta| + |\overline{\beta_2}| \leq |\gamma| + |\overline{\beta_1}| + |\overline{\beta_2}| \leq |\overline{\beta}| + |\overline{\beta_1}| + |\overline{\beta_2}| \leq 3m.$$

It follows that $d(uv) \leq 3m$ which implies that $(uv)\theta \in Y$ and, by the definition of k, we can write $uv = u_1v_1$ where $|v_1| \leq k$.

This completes the proof of the claim. □

Returning to the proof of Proposition 7.6, let W be the set of words of $\mathcal{L}(A,U)$ that have length no greater than k. This set is finite and as a consequence so is the set $W\theta$. It now follows from the claim that $U^2 \cap T \subseteq U(W\theta) \cap T$ which, by condition (i), is a finite set. □

7.3 Generating subsemigroups using boundaries

In this section we will prove the first main result of this chapter:

Theorem 7.1 *If S is a finitely generated semigroup and T is a subsemigroup of S with finite boundary then T is finitely generated.*

The right (or left) boundary of a subsemigroup T of a semigroup S may be used to construct a generating set for T.

Proposition 7.7. *Let S be a finitely generated semigroup and let T be a subsemigroup of S. Let A be a finite generating set for S. Then either of the sets below generates T:*

$$X_\rho = \mathcal{B}_r(A,T)U^1 \cap T, \qquad X_\lambda = U^1\mathcal{B}_l(A,T) \cap T.$$

Proof. Let $t \in T$ be arbitrary. Write $t = a_1 \dots a_k$ where $a_i \in A$ for $1 \leq i \leq k$. Let m be the smallest subscript such that $\beta_1 = a_1 \dots a_m$ belongs to T, let $\gamma_1 = a_{m+1} \dots a_k$ and note that $\beta_1 \in \mathcal{B}_r(A, T)$. If $\gamma_1 \in U$ or is empty then stop. Otherwise repeat the same process on the word $a_{m+1} \dots a_k$ writing it as $\beta_2\gamma_2$ where β_2 is the shortest prefix that belongs to $\mathcal{L}(A, T)$ so that $\beta_2 \in \mathcal{B}_r(A, T)$. Continuing in this way, in a finite number of steps, we can write $t = \beta_1 \dots \beta_{m-1}\beta_m\gamma_m$ where $\beta_1, \dots, \beta_m \in \mathcal{B}_r(A, T)$ and $\gamma_m \in U^1$. The elements $\beta_1, \dots, \beta_{m-1}, \beta_m\gamma_m$ all belong to X_ρ and, since t was arbitrary, it follows that X_ρ generates T. The fact that X_λ generates T follows from a dual argument. □

Proof of Theorem 7.1. If $\mathcal{B}_l(A, T)$ and $\mathcal{B}_r(A, T)$ are finite then X_ρ is finite, by Proposition 7.6 (ii), and generates T, by Proposition 7.7. □

We can use this result to give an upper bound for the rank of T based on the size and depth of the boundary.

Corollary 7.8. *Let S be a semigroup generated by A and let T be a subsemigroup of S. Then*

$$\operatorname{rank}(T) \leq d(\mathcal{B}_r(A, T))|\mathcal{B}_r(A, T)||\mathcal{B}_l(A, T)|.$$

Proof. Let $w \in \mathcal{L}(A, X_\rho)$ and let $Y \subseteq \mathcal{L}(A, \mathcal{B}_r(A, T))$ be a fixed set of shortest length word representatives of the elements of $\mathcal{B}_r(A, T)$. Then $w = vu$ where $v \in Y$ and $u \in \mathcal{L}(A, U^1)$. Since $u \in \mathcal{L}(A, U^1)$ and $vu \in \mathcal{L}(A, T)$ we can write $vu \equiv v'\beta$ where $\beta \in \mathcal{L}(\mathcal{B}_l(A, T))$ and v' is a prefix of v. Since w was arbitrary it follows that every word in $\mathcal{L}(A, X_\rho)$ is equal to a word w_1w_2 where w_1 is a prefix, possibly the empty word, of a word from Y and $w_2 \in \mathcal{L}(A, \mathcal{B}_l(A, T)) \cap \mathcal{R}$. The result follows since there are at most $d(\mathcal{B}_r(A, T))|\mathcal{B}_r(A, T)|$ choices for w_1 and $|\mathcal{B}_l(A, T)|$ choices for w_2. □

The above bound is attained infinitely often as the following example shows.

Example 7.9. Let $S = \mathbb{Z}$ the infinite cyclic group written additively so that it is generated, as a semigroup, by $\{-1, 1\}$. Let $m \in \mathbb{N}$ and let $T = \{n \in \mathbb{Z} : n \geq m\}$. Then T is a subsemigroup of S, T does not have finite Rees index in S, and

$$\operatorname{rank}(T) = m = d(\mathcal{B}_r(A, T))|\mathcal{B}_r(A, T)||\mathcal{B}_l(A, T)|.$$

The generating set given in the Proposition 7.7 is best possible in general as the following example demonstrates.

Example 7.10. Let A and B be finite alphabets and let S be the semigroup defined by the presentation

$$\langle A, B, 0 \mid ab = ba = 0,\ a0 = 0a = b0 = 0b = 00 = 0,\ a \in A,\ b \in B\rangle.$$

The semigroup S is generated by the set $A \cup B$. Let $T = \{w \in A^+ \mid |w| \geq k\} \cup \{0\}$ which is a subsemigroup of T and has infinite Rees index in T. Then

$$\mathcal{B}_r(A \cup B, T) = \{w \in A^+ \mid |w| = k\} \cup \{0\}$$

and

$$U^1 = \{\epsilon\} \cup B^+ \cup \{a \in A^+ \mid |w| < k\}.$$

Therefore

$$X_\rho = \mathcal{B}_r(A \cup B, T)U^1 \cap T = \{w \in A^+ \mid k \leq |w| \leq 2k - 1\} \cup \{0\}.$$

Since $T \setminus \{0\}$ is a subsemigroup of the free semigroup A^+ it follows from [57, Chapter 7] that the unique minimal generating set for T is $T \setminus T^2 \cup \{0\} = X_\rho$.

Note that if only the right (or left) boundary is finite then T need not inherit the property of being finitely generated as the following example demonstrates.

Example 7.11. Let $F = A^+$, the free semigroup over the alphabet A, where $A = \{a, b\}$. Let R be the subsemigroup of all words that begin with the letter a. Then $\mathcal{B}_r(A, T) = \{a\}$ which is finite but R is not finitely generated since all the elements ab^i where $i \in \mathbb{N}$ must be included in any generating set.

7.4 Presentations

Preliminaries: definitions and notation

A semigroup presentation is a pair $\mathfrak{P} = \langle A|\mathfrak{R}\rangle$ where A is a an alphabet and $\mathfrak{R} \subseteq A^+ \times A^+$ is a set of pairs of words. An element (u, v) of $\mathfrak{R}$ is called a *relation* and is usually written $u = v$. We say that S is the *semigroup defined by the presentation* $\mathfrak{P}$ if $S \cong A^+/\eta$ where η is the smallest congruence on A^+ containing $\mathfrak{R}$. We may think of S as the largest semigroup generated by the set A which satisfies all the relations of $\mathfrak{R}$. We say that a semigroup S is finitely presented if it can be defined by $\langle A|R\rangle$ where A and R are both finite. For example the free semigroup on the alphabet $\{a, b\}$ is given by the presentation $\langle a, b|\rangle$ and hence is finitely presented. At the other extreme, every finite semigroup is finitely presented: by including the entire multiplication table in the set of relations if necessary. Note that not every semigroup is finitely presented: consider the semigroup defined by the presentation $\langle a, b \mid ab^i a = aba,\ (i \in \mathbb{N})\rangle$ for example. If S is a finitely presented semigroup and T is a finitely generated subsemigroup of S then T need not be finitely presented, as the following example demonstrates.

Example 7.12. Let $S = \{a, b, c\}^+$, a free semigroup of rank 3, and let T be the subsemigroup of S generated by the set $X = \{ba, ba^2, a^3, ac, a^2c\}$. Let $\xi_1 = ba$, $\xi_2 = ba^2$, $\xi_3 = a^3$, $\xi_4 = ac$, and $\xi_5 = a^2c$. The relations

$$\xi_2 {\xi_3}^n \xi_4 = \xi_1 {\xi_3}^n \xi_5, \quad (n \in \mathbb{N})$$

all hold in T and they are irreducible in the sense that no non-trivial relation may be applied to either side of the relation. It follows that T is not finitely presented.

We say that the word $w \in A^+$ *represents the element* $s \in S$ if $s = w/\eta$. As in §7.2 given two words $w, v \in A^+$ we write $w = v$ if w and v represent the same element of S and write $w \equiv v$ if w and v are identical as words. Also, given an element $s \in S$ and a word $w \in A^+$ we write $w = s$ when $w/\eta = s$ in S.

We say that w is obtained from v by one application of a relation from $\mathfrak{R}$ if there exist $\alpha, \beta \in A^*$ and $(x = y) \in \mathfrak{R}$ such that $w \equiv \alpha x \beta$ and $v \equiv \alpha y \beta$. We say that the relation $w = v$ is a consequence of the relations $\mathfrak{R}$ (or of the presentation $\mathfrak{P}$) if there is a finite sequence of words $(\alpha_1, \ldots, \alpha_m)$ such that $w \equiv \alpha_1$, $v \equiv \alpha_m$ and, for all k, α_{k+1} is obtained from α_k by one application of a relation from $\mathfrak{R}$. We now state a basic result that will be used frequently in what follows.

Proposition 7.13. *Let $\mathfrak{P} = \langle A | \mathfrak{R}\rangle$ be a semigroup presentation, let $S = A^+/\eta$ be the semigroup defined by it, and let $\alpha, \beta \in A^+$ be any two words. Then the relation $\alpha = \beta$ holds in S if and only if it is a consequence of $\mathfrak{P}$.*

The Reidemeister–Schreier theorem for semigroups

Before embarking on the proof of the second main result of this chapter we need to give some background on the Reidemeister–Schreier theorem for semigroups described in [17]. This theorem is the semigroup analogue of Theorem 2.6 of [72]. It gives a general method for finding presentations for subsemigroups of a given finitely presented semigroup.

Begin with a semigroup S defined by a presentation $\langle A|R\rangle$. We will find a presentation $\langle B|Q\rangle$ for a subsemigroup T of S where every element of B corresponds to a word in A^+. More precisely, let T be the subsemigroup of S generated by $X\theta$ where $X = \{\xi_i : i \in I\}$ with the ξ_i being words from A^+. Introduce a new alphabet $B = \{b_i : i \in I\}$ in 1–1 correspondence with the set X. Define $\psi : B^+ \to A^+$ by extending the map $\psi(b_i) = \xi_i$, $i \in I$. We call ψ the *representation mapping* since given any $w \in B^+$ the word $w\psi \in A^+$ represents an element of T.

We say that a mapping $\phi : \mathcal{L}(A,T) \to B^+$ is a *rewriting mapping* if it satisfies $w\phi\psi = w$ in S for all $w \in \mathcal{L}(A,T)$. Note that this is not the same as saying that $w\phi\psi \equiv w$ in A^+. The map ϕ may be thought of as rewriting every $w \in \mathcal{L}(A,T)$ as a product of the given generators for T. Such a representation mapping always exists since for each word w in $\mathcal{L}(A,T)$ there is a set of products in B^+ that equal w and we can choose one of them.

Theorem 7.14. *[17, Theorem 2.1] Let S be the semigroup defined by a presentation $\langle A|R\rangle$, and let T be the subsemigroup of S generated by $X = \{\xi_i : i \in I\} \subseteq A^+$. Introduce a new alphabet $B = \{b_i : i \in I\}$, and let ψ and ϕ be the representation mapping and the rewriting mapping. Then T is defined by generators B and relations:*

$$b_i = \xi_i\phi, \; i \in I \tag{7.1}$$

$$(w_1w_2)\phi = (w_1\phi)(w_2\phi) \tag{7.2}$$

$$(w_3uw_4)\phi = (w_3vw_4)\phi \tag{7.3}$$

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 $w_3uw_4 \in \mathcal{L}(A,T)$.

If T is finitely generated then the set of relations (7.1) will be finite. The sets of relations (7.2) and (7.3) will not, in general, be finite. A pictorial representation of the theorem is given in Figure 7.3.

One problem with Theorem 7.14 is that it asserts the existence of a rewriting mapping without actually constructing one. Also, the presentation it produces is necessarily infinite even if T is finitely presented. The usefulness of the theorem is that it may be used as a tool for finding finite presentations for subsemigroups in various special circumstances. In certain situations it is possible to pin down a specific rewriting mapping and then build a presentation with finitely many relations from which all of the relations given in Theorem 7.14 may be deduced. This approach has been used successfully in the past. For example, in the special case where T is a two-sided ideal with finite Rees index a finite presentation may be explicitly written down for T (see [17, Theorem 4.1]).

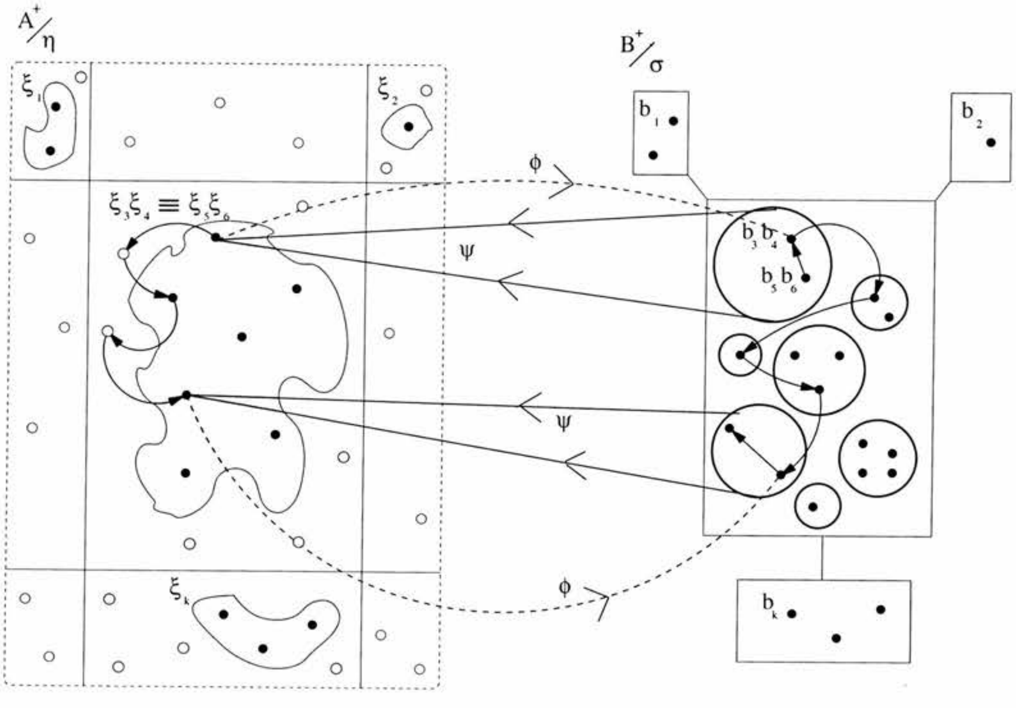


Figure 7.3: A diagram representing the Reidemeister–Schreier rewriting process for semigroups. On the left the semigroup S is represented as the factor semigroup A^+/η and on the right the subsemigroup T is represented as the factor semigroup B^+/σ. Here η is the smallest congruence on A^+ containing R and σ is the smallest congruence on B^+ containing Q, where Q is the set of relations given in Theorem 7.14. The "boxes" represent the η- and σ- classes of A^+ and B^+ respectively. The subsemigroup T is generated by the words $W = \{\xi_i : i \in I\}$ and the shaded areas on the left correspond to words in the subsemigroup of the free semigroup A^+ generated by W. Note that $\langle W \rangle \subseteq \mathcal{L}(A,T)$ in A^+ and $\mathcal{L}(A,T)$ is just the union of the η-classes of A^+ that $\langle W \rangle$ intersects (in the picture this is the union of the boxes that contain some shading). The words in $\langle W \rangle$ are not quite in one to one correspondence with the words in B^+ since the map ψ may not be injective (in the picture $\xi_3\xi_4 \equiv \xi_5\xi_6$ in A^+ while b_3b_4 and b_5b_6 are distinct in B^+). Carrying out deductions on words of B^+ using the relations from Q we first use relations (7.2) to move to a word in the image of $\mathcal{L}(A,T)$ under ϕ. This may be thought of as a "bracket rearranging process". We then mirror a sequence of deductions in A^+, using relations from R, in the image of A^+ under ϕ, using the relations (7.3). Finally we again use relations (7.2) to move to the appropriate word within the target ψ^{-1}-class of B^+. An example of such a sequence of deductions is given by the directed path indicated in the diagram.

Proof of Theorem 7.2: bars and hats

We now begin the task of proving Theorem 7.2. We will use the Reidemeister–Schreier theorem described above to construct a presentation for T from a given presentation of S. This presentation will be of the form $\langle B|\mathcal{Q}\rangle$ where B is finite and $\mathcal{Q}$ is infinite. We then go on to prove that there exists a finite set of relations $\mathcal{D}$ that all hold in T with the property that every relation of $\mathcal{Q}$ is a consequence of the relations $\mathcal{D}$. It will follow that T is defined by $\langle B|\mathcal{D}\rangle$ where B and $\mathcal{D}$ are both finite.

Let S be the semigroup defined by the presentation $\mathfrak{P} = \langle A|\mathfrak{R}\rangle$ and let η be the smallest congruence on A^+ containing $\mathfrak{R}$. Let T be a subsemigroup of S with a finite boundary.

As before, fix a transversal $\mathcal{R}$ of the η-classes of A^+ chosen so that every $w \in \mathcal{R}$ is a word of shortest length in its η-class. For every $w \in A^+$ define $\overline{w} = (w/\eta) \cap \mathcal{R}$: the fixed shortest length word in $\mathcal{R}$ that equals w in S.

Define $\mathcal{PI}(A,T) \subseteq \mathcal{L}(A,T)$ to be the set of words w such that every prefix of w (with the exception of w itself) belongs to $\mathcal{L}(A,U^1)$. We call $\mathcal{PI}(A,T)/\eta$ the *strict right boundary of* T *in* S. Note that that strict right boundary of T in S is a subset of the right boundary of T. In Section 7.3 we found a generating set for T by multiplying the elements of the right boundary on the right by elements of U. In fact, if we just take the strict right boundary and multiply on the right by elements of U we obtain a generating set. This is the generating set with respect to which we will write a presentation for T. Let $\mathcal{SB}_r(A,T) = \mathcal{PI}(A,T)/\eta$ and define

$$X_\rho = \mathcal{SB}_r(A,T)U^1 \cap T$$

which is a generating set for T by exactly the same argument as in the proof of Lemma 7.7.

We will now partition the elements of U into classes depending on how they interact with T.

Definition 7.15. Choose a symbol $0 \notin S$ and for each $u \in U$ define $f_u, g_u : T \cup \{0\} \to T \cup \{0\}$ by:

$$xf_u = \begin{cases} xu & \text{if } x \in T, xu \in T \\ 0 & \text{otherwise,} \end{cases} \qquad xg_u = \begin{cases} ux & \text{if } x \in T, ux \in T \\ 0 & \text{otherwise.} \end{cases}$$

Then, given a subset X of T and given $u, v \in U$ we write $u \sim_X v$ if and only if $f_u \restriction_X = f_v \restriction_X$ and $g_u \restriction_X = g_v \restriction_X$. (Here $f_u \restriction_X$ denotes the map f_u restricted to the set X.)

The relation $\sim_X$ so defined is clearly an equivalence relation. Moreover, if X is finite then, by Proposition 7.6, $\sim_X$ has finitely many equivalence classes. Indeed, we have $f_u \restriction_X: X \to UX \cup \{0\}$ and $g_u \restriction_X: X \to XU \cup \{0\}$ and, by Proposition 7.6, the ranges of both of these functions are finite. In particular, since the generating set X_ρ is finite, the subset U has finitely many $\sim_{X_\rho}$-classes.

Given $w, u \in \mathcal{L}(A, U^1)$ we write $w \approx_X u$ if $w/\eta \sim_X u/\eta$. Again, this is an equivalence relation. Let $\Sigma \subseteq \mathcal{L}(A, U^1)$ be a set of smallest length word representatives of the $\approx_{X_\rho}$ classes. Clearly $\epsilon \in \Sigma$.

We define the operation *hat* ($w \mapsto \widehat{w}$) on the words of $\mathcal{L}(A, U^1)$ by $\{\widehat{w}\} = (w/\eta) \cap \Sigma$ and $\widehat{\epsilon} = \epsilon$. Note that any word in A^* may be barred but only words in $\mathcal{L}(A, U^1)$ may be hatted.

The following lemma summarises several properties of the hat operation. Its proof is an immediate consequence of the above definitions and discussion.

Lemma 7.16. *The following properties hold:*

(i) For $u \in \mathcal{L}(A, U)$ we have $|\widehat{u}| \leq |u|$.

(ii) The set Σ is finite.

(iii) If $\gamma \in \mathcal{L}(A, X_\rho), u \in \mathcal{L}(A, U)$ and $u\gamma \in \mathcal{L}(A, T)$ then $u\gamma = \widehat{u}\gamma$.

(iv) If $\gamma \in \mathcal{L}(A, X_\rho), u \in \mathcal{L}(A, U)$ and $\gamma u \in \mathcal{L}(A, T)$ then $\gamma u = \gamma\widehat{u}$.

Proof (continued): representation and rewriting mappings

In this section we find a presentation for T. It will have infinitely many relations and the rest of the proof will be devoted to reducing this infinite set to a finite one.

The generating set X_ρ can be represented by the following set of words in A^+:

$$\{\overline{v}u : v \in \mathcal{PI}(A, T), u \in \Sigma, vu \in \mathcal{L}(A, T)\}.$$

We construct a new alphabet B in one-one correspondence with these generating words:

$$B = \{b_{\overline{v},u} : v \in \mathcal{PI}(A, T), u \in \Sigma, vu \in \mathcal{L}(A, T)\}.$$

This set is finite since A is finite and T has a finite boundary in S.

Let $\psi : B^+ \to A^+$ be the unique homomorphism extending $b_{\overline{v},u} \mapsto \overline{v}u$ and, using the terminology introduced earlier in the section, call this map the *representation mapping*. It has the property that for every $w \in B^+$, the words w and $w\psi \in A^+$ represent the same element of S (and, of course, of T).

Now define a map $\phi : \mathcal{L}(A,T) \to B^+$ as follows. For $w \in \mathcal{L}(A,T)$ write $w \equiv \alpha\beta$ where $\alpha \in A^+, \beta \in A^*$ and α is the shortest prefix of w belonging to $\mathcal{L}(A,T)$: so α is the unique prefix of w that belongs to $\mathcal{PI}(A,T)$. Then ϕ is defined inductively by:

$$w\phi = \begin{cases} b_{\overline{\alpha},\widehat{\beta}} & \text{if } \beta \notin \mathcal{L}(A,T) \\ b_{\overline{\alpha},\epsilon}(\beta\phi) & \text{if } \beta \in \mathcal{L}(A,T). \end{cases}$$

It is easy to see that for every $w \in \mathcal{L}(A,T)$ the relation $w\phi\psi = w$ holds in S. (But we usually have $w\phi\psi \not\equiv w$.) Using the terminology introduced earlier in the section, the map ϕ is a *rewriting mapping*.

It now follows from Theorem 7.14 that the semigroup T is defined by the presentation with generators B and relations:

$$b_{\overline{v},u} = (\overline{v}u)\phi \tag{7.4}$$

$$(w_1w_2)\phi = (w_1\phi)(w_2\phi) \tag{7.5}$$

$$(w_3xw_4)\phi = (w_3yw_4)\phi \tag{7.6}$$

where $v \in \mathcal{PI}(A,T)$, $u \in \Sigma$, $vu \in \mathcal{L}(A,T)$, $w_1, w_2 \in \mathcal{L}(A,T)$, $w_3, w_4 \in A^*$, $(x = y) \in \mathfrak{R}$, $w_3xw_4 \in \mathcal{L}(A,T)$.

The set of relations (7.4) is finite since B is finite. The remainder of the proof is concerned with proving that the relations (7.5) and (7.6) are all consequences of a fixed finite set of relations $\mathcal{D}$ that we define below.

Before we do that, we state a lemma which gives a canonical decomposition of words from $\mathcal{L}(A,T)$ that is compatible with the operation of ϕ. The proof is an immediate consequence of the definition of ϕ.

Lemma 7.17. *Let $w \in \mathcal{L}(A,T)$ be arbitrary. The word w can be written uniquely as*

$$w \equiv \alpha_1 \dots \alpha_{k-1}\alpha_k\alpha_{k+1}$$

where $k \geq 1$, $\alpha_1, \dots, \alpha_k \in \mathcal{PI}(A,T), \alpha_{k+1} \in \mathcal{L}(A,U^1)$ and $\alpha_k\alpha_{k+1} \in \mathcal{L}(A,T)$. When applying the rewriting mapping we obtain:

$$w\phi \equiv (\alpha_1\phi)\dots(\alpha_{k-1}\phi)(\alpha_k\alpha_{k+1})\phi \equiv b_{\overline{\alpha_1},\epsilon} \dots b_{\overline{\alpha_{k-1}},\epsilon} b_{\overline{\alpha_k},\widehat{\alpha_{k+1}}}.$$

We call the words $\alpha_1, \dots, \alpha_{k-1}, \alpha_k\alpha_{k+1}$ the *principal factors* of w and when we write $w \in \mathcal{L}(A,T)$ as $\alpha_1 \dots \alpha_{k-1}\alpha_k\alpha_{k+1}$ we say that it has been *decomposed* into principal factors.

Proof (continued): a finite set $\mathcal{D}$ of relations

We use the fact that T has a finite boundary in S and that, by Proposition 7.6, $U^2 \cap T$ is finite to define the following four numbers:

(i) $M_{\mathcal{B}} = \max\{|\overline{\gamma\delta}| : \gamma \in \mathcal{PI}(A,T), \delta \in \mathcal{L}(A,U^1), \gamma\delta \in \mathcal{L}(A,T)\}$ (well defined by Proposition 7.7);

(ii) $M_{\Sigma} = \max\{|\sigma| : \sigma \in \Sigma\} = \max\{|\widehat{u}| : u \in \mathcal{L}(A,U^1)\}$ (by Lemma 7.16 part (ii));

(iii) $M_{UU} = \max\{|\overline{uv}| : u,v \in \mathcal{L}(A,U^1), uv \in \mathcal{L}(A,T)\}$ (by Proposition 7.6);

(iv) $M_{\mathfrak{R}} = \max\{|uv| : (u = v) \in \mathfrak{R}\}$ (well defined since $\mathfrak{R}$ is finite).

Let $\mathcal{D}$ be the set of all relations in the alphabet B which hold in T and have length that does not exceed

$$N = 4(\max\{M_{\mathcal{B}}, M_{\Sigma}, M_{UU}, M_{\mathfrak{R}}\} + 1).$$

In other words $\mathcal{D} = \{(u,v) \in B^+ \times B^+ : |uv| \leq N, u\psi = v\psi \text{ holds in } S\}$. The rest of this section will be spent proving the following theorem.

Theorem 7.18. *The presentation $\langle B|\mathcal{D}\rangle$ defines T.*

Proof (continued): three technical lemmas

We now present three key lemmas that are used to prove Theorem 7.18.

Lemma 7.19. *The relations $(uv)\phi = (\overline{uv})\phi$ where $u,v \in \mathcal{L}(A,U^1)$ and $uv \in \mathcal{L}(A,T)$ are consequences of $\mathcal{D}$.*

Proof. We proceed by induction on the length of the word uv. When $|uv| \leq 2$ the relation $(uv)\phi = (\overline{uv})\phi$ is in $\mathcal{D}$ since

$$|(uv)\phi(\overline{uv})\phi| = |(uv)\phi| + |(\overline{uv})\phi| \leq |uv| + |\overline{uv}| \leq 2 + 2 = 4.$$

Now let $u,v \in \mathcal{L}(A,U^1)$ where $uv \in \mathcal{L}(A,T)$ and suppose that the result holds for all $u_1, v_1 \in \mathcal{L}(A,U^1)$ with $u_1v_1 \in \mathcal{L}(A,T)$ and $|u_1v_1| < |uv|$. There are two cases to consider depending on whether or not u has a prefix in $\mathcal{L}(A,T)$.

Case 1: *u has a prefix in $\mathcal{L}(A,T)$.* Write $u \equiv u'u''$ where u' is the shortest such prefix. Since T is a subsemigroup, $u \in \mathcal{L}(A,U^1)$ and $u' \in \mathcal{L}(A,T)$ it follows that $u'' \notin \mathcal{L}(A,T)$. The case now splits into two subcases.

Case 1.1: $u''v \notin \mathcal{L}(A,T)$. In this case $(uv)\phi$ is a single letter and $(uv)\phi = (\overline{uv})\phi$ belongs to $\mathcal{D}$ since

$$|(uv)\phi(\overline{uv})\phi| = 1 + |(\overline{uv})\phi| \leq 1 + |\overline{uv}| \leq 1 + M_{UU}.$$

Case 1.2: $u''v \in \mathcal{L}(A,T)$. In this case since $u'', v \in \mathcal{L}(A,U^1)$, $u''v \in \mathcal{L}(A,T)$ and $|u''v| < |uv|$ we can apply induction giving

$$\begin{array}{rcll} (uv)\phi & \equiv & (u')\phi(u''v)\phi & \text{(by Lemma 7.17)} \\ & = & (u')\phi(\overline{u''v})\phi & \text{(induction)} \\ & = & (\overline{uv})\phi & \text{(in } \mathcal{D}). \end{array}$$

In the last step the relation $(u')\phi(\overline{u''v})\phi = (\overline{uv})\phi$ is in $\mathcal{D}$ since

$$|(u')\phi(\overline{u''v})\phi\ (\overline{uv})\phi| = 1 + |(\overline{u''v})\phi| + |(\overline{uv})\phi| \leq 1 + |\overline{u''v}| + |\overline{uv}| \leq 1 + 2M_{UU}.$$

Case 2: *u has no prefix in $\mathcal{L}(A,T)$.* First decompose $uv \equiv u\beta_1 \ldots \beta_b\beta_{b+1}$ where the principal factors are $u\beta_1, \beta_2, \ldots, \beta_{b-1}, \beta_b\beta_{b+1}$. We follow the convention that β_1 always exists and β_{b+1} may be the empty word. This case now splits into two subcases.

Case 2.1: $b = 1$. In this case $(uv)\phi$ is a single letter and $(uv)\phi = (\overline{uv})\phi$ is in $\mathcal{D}$ since it has length $|(uv)\phi(\overline{uv})\phi| \leq 1 + |\overline{uv}| \leq 1 + M_{UU}$.

Case 2.2: $b \geq 2$. First note that since $\beta_b\beta_{b+1} \in \mathcal{L}(A,T)$ and $v \notin \mathcal{L}(A,T)$ it follows that $\beta_1 \ldots \beta_{b-1} \notin \mathcal{L}(A,T)$. Then we have:

$$\begin{array}{rcll} (uv)\phi & \equiv & (u\beta_1 \ldots \beta_{b-1}\beta_b\beta_{b+1})\phi & \\ & \equiv & (u\beta_1 \ldots \beta_{b-1})\phi(\beta_b\beta_{b+1})\phi & \text{(by Lemma 7.17)} \\ & = & (\overline{u\beta_1 \ldots \beta_{b-1}})\phi(\beta_b\beta_{b+1})\phi & \text{(induction)} \\ & = & (\overline{uv})\phi & \text{(in } \mathcal{D}). \end{array}$$

In the last step the relation $(\overline{u\beta_1 \ldots \beta_{b-1}})\phi(\beta_b\beta_{b+1})\phi = (\overline{uv})\phi$ is in $\mathcal{D}$ since

$$\begin{array}{rcl} |(\overline{u\beta_1 \ldots \beta_{b-1}})\phi(\beta_b\beta_{b+1})\phi(\overline{uv})\phi| & = & |(\overline{u\beta_1 \ldots \beta_{b-1}})\phi| + |(\beta_b\beta_{b+1})\phi| + |(\overline{uv})\phi| \\ & \leq & |(\overline{u\beta_1 \ldots \beta_{b-1}})| + 1 + |(\overline{uv})| \leq 1 + 2M_{UU} \end{array}$$

as required.

□

Lemma 7.20. *The relations* $(\beta\gamma\delta)\phi = (\widehat{\beta}\ \overline{\gamma\delta})\phi$ *where* $\beta \in \mathcal{L}(A, U^1), \gamma \in \mathcal{PI}(A,T), \delta \in \mathcal{L}(A, U^1), \gamma\delta \in \mathcal{L}(A,T)$ *and* $\beta\gamma\delta \in \mathcal{L}(A,T)$ *are consequences of* $\mathcal{D}$.

Proof. We proceed by induction on the length of the word $|\beta\gamma\delta|$. When $|\beta\gamma\delta| \leq 3$ the relation $(\beta\gamma\delta)\phi = (\widehat{\beta}\ \overline{\gamma\delta})\phi$ is in $\mathcal{D}$ since

$$|(\beta\gamma\delta)\phi(\widehat{\beta}\ \overline{\gamma\delta})\phi| \leq |\beta\gamma\delta| + |\widehat{\beta}| + |\overline{\gamma\delta}| \leq 3 + M_\Sigma + M_\mathcal{B}.$$

Now let $\beta \in \mathcal{L}(A, U^1), \gamma \in \mathcal{PI}(A,T), \delta \in \mathcal{L}(A, U^1)$ and $\gamma\delta \in \mathcal{L}(A,T)$ where $\beta\gamma\delta \in \mathcal{L}(A,T)$, and suppose that the result holds for $\beta_1, \gamma_1, \delta_1$ satisfying the analogous conditions and with $|\beta_1\gamma_1\delta_1| < |\beta\gamma\delta|$. First observe that if β is empty then the relation becomes $(\gamma\delta)\phi = (\overline{\gamma\delta})\phi$ which has length $|(\gamma\delta)\phi(\overline{\gamma\delta})\phi| \leq 1 + M_\mathcal{B}$ and so belongs to $\mathcal{D}$. When β is not empty there are two cases to consider depending on whether or not β has a prefix in $\mathcal{L}(A,T)$.

Case 1: *β has a prefix in $\mathcal{L}(A,T)$.* Let β' be the shortest such prefix and write $\beta \equiv \beta'\beta''$. Note that since $\beta \in \mathcal{L}(A, U^1)$, $\beta' \in \mathcal{L}(A,T)$ and T is a subsemigroup of S it follows that $\beta'' \in \mathcal{L}(A, U^1)$. This case now splits into two subcases depending on whether or not $\beta''\gamma\delta \in \mathcal{L}(A,T)$.

Case 1.1: $\beta''\gamma\delta \notin \mathcal{L}(A,T)$. In this case $(\beta\gamma\delta)\phi$ is a single letter and $(\beta\gamma\delta)\phi = (\widehat{\beta}\ \overline{\gamma\delta})\phi$ is in $\mathcal{D}$ since

$$|(\beta\gamma\delta)\phi(\widehat{\beta}\ \overline{\gamma\delta})\phi| = 1 + |(\widehat{\beta}\ \overline{\gamma\delta})\phi| \leq 1 + |\widehat{\beta}\ \overline{\gamma\delta}| = 1 + |\widehat{\beta}| + |\overline{\gamma\delta}| \leq 1 + M_\Sigma + M_\mathcal{B}.$$

Case 1.2: $\beta''\gamma\delta \in \mathcal{L}(A,T)$. In this case we have:

$$\begin{aligned} (\beta\gamma\delta)\phi &\equiv (\beta'\beta''\gamma\delta)\phi \\ &\equiv (\beta')\phi(\beta''\gamma\delta)\phi \quad \text{(by Lemma 7.17)} \\ &= (\beta')\phi(\widehat{\beta''}\ \overline{\gamma\delta})\phi \quad \text{(induction)} \\ &= (\widehat{\beta}\ \overline{\gamma\delta})\phi \quad \text{(in } \mathcal{D}\text{)}. \end{aligned}$$

In the last step the relation $(\beta')\phi(\widehat{\beta''}\overline{\gamma\delta})\phi = (\widehat{\beta}\ \overline{\gamma\delta})\phi$ is in $\mathcal{D}$ since

$$\begin{aligned} |(\beta')\phi(\widehat{\beta''}\ \overline{\gamma\delta})\phi(\widehat{\beta}\ \overline{\gamma\delta})\phi| &= 1 + |(\widehat{\beta''}\ \overline{\gamma\delta})\phi| + |(\widehat{\beta}\ \overline{\gamma\delta})\phi| \leq 1 + |\widehat{\beta''}\ \overline{\gamma\delta}| + |\widehat{\beta}\ \overline{\gamma\delta}| \\ &= 1 + |\widehat{\beta''}| + |\overline{\gamma\delta}| + |\widehat{\beta}| + |\overline{\gamma\delta}| \leq 1 + 2M_\Sigma + 2M_\mathcal{B}. \end{aligned}$$

Case 2: *β has no prefix in $\mathcal{L}(A,T)$.* In this case we decompose:

$$\beta\gamma\delta \equiv \beta\gamma_1 \dots \gamma_{c+1}\delta_1 \dots \delta_d\delta_{d+1}$$

where the principal factors are $\beta\gamma_1, \gamma_2, \dots, \gamma_c, \gamma_{c+1}\delta_1, \delta_2, \dots, \delta_{d-1}, \delta_d\delta_{d+1}$.

A few words of explanation are in order here. As usual, we think of the principal factors as being obtained by reading the word $\beta\gamma\delta$ from left to right, and writing successive prefixes that belong to $\mathcal{PI}(A,T)$, as long as the remaining suffix is in $\mathcal{L}(A,T)$. Thus, $\beta\gamma_1$ is the first such prefix, provided it is also a prefix of $\beta\gamma$. If the first such prefix is longer than $\beta\gamma$ we take $c=0$ and $\gamma_{c+1} \equiv \gamma_1 \equiv \gamma$. Also, γ_c is the last of these prefixes which ends inside γ, and γ_{c+1} is the rest of γ. Of course, it may happen that γ_c ends at the last letter of γ, in which case we take $\gamma_{c+1} \equiv \epsilon$. Furthermore, in this case, $\gamma\delta$ is the final principal factor since $\delta \notin \mathcal{L}(A,T)$ and so we take $d=0$ and $\delta_{d+1} \equiv \delta_1 \equiv \delta$.

This case now splits into two subcases.

Case 2.1: $d \geq 2$. In this case, by the definition of ϕ, we have

$$(\beta\gamma\delta)\phi \equiv (\beta\gamma\delta_1 \dots \delta_{d-1})\phi(\delta_d\delta_{d+1})\phi.$$

This subcase now splits into two subcases depending on whether or not $\gamma\delta_1 \dots \delta_{d-1} \in \mathcal{L}(A,T)$.

Case 2.1.1: $\gamma\delta_1 \dots \delta_{d-1} \notin \mathcal{L}(A,T)$. Then since $\beta \notin \mathcal{L}(A,T)$ we can apply the previous lemma to give:

$$\begin{aligned} (\beta\gamma\delta)\phi &\equiv (\beta\gamma\delta_1 \dots \delta_{d-1})\phi(\delta_d\delta_{d+1})\phi \\ &= (\overline{\beta\gamma\delta_1 \dots \delta_{d-1}})\phi(\delta_d\delta_{d+1})\phi \quad \text{(by Lemma 7.19)} \\ &= (\widehat{\beta}\,\overline{\gamma\delta})\phi \quad \text{(in } \mathcal{D}\text{)}. \end{aligned}$$

In the last step the relation $(\overline{\beta\gamma\delta_1 \dots \delta_{d-1}})\phi(\delta_d\delta_{d+1})\phi = (\widehat{\beta}\,\overline{\gamma\delta})\phi$ is in $\mathcal{D}$ since

$$\begin{aligned} |(\overline{\beta\gamma\delta_1 \dots \delta_{d-1}})\phi(\delta_d\delta_{d+1})\phi(\widehat{\beta}\,\overline{\gamma\delta})\phi| &= |(\overline{\beta\gamma\delta_1 \dots \delta_{d-1}})\phi| + 1 + |(\widehat{\beta}\,\overline{\gamma\delta})\phi| \\ &\leq |\overline{\beta\gamma\delta_1 \dots \delta_{d-1}}| + 1 + |\widehat{\beta}| + |\overline{\gamma\delta}| \\ &\leq M_{UU} + 1 + M_\Sigma + M_\mathcal{B}. \end{aligned}$$

Case 2.1.2: $\gamma\delta_1 \dots \delta_{d-1} \in \mathcal{L}(A,T)$. Then, since $\delta_d\delta_{d+1} \in \mathcal{L}(A,T)$, $\delta \in \mathcal{L}(A,U^1)$ and T is a subsemigroup of S, we have $\delta_1 \dots \delta_{d-1} \notin \mathcal{L}(A,T)$ and so $\gamma\delta_1 \dots \delta_{d-1} \in$

$\mathcal{L}(A, X_\rho)$ and we can apply induction giving:

$$\begin{array}{rcll} (\beta\gamma\delta)\phi & \equiv & (\beta\gamma\delta_1\ldots\delta_{d-1})\phi(\delta_d\delta_{d+1})\phi & \text{(by Lemma 7.17)} \\ & = & (\widehat{\beta}\,\overline{\gamma\delta_1\ldots\delta_{d-1}})\phi(\delta_d\delta_{d+1})\phi & \text{(induction)} \\ & = & (\widehat{\beta}\,\overline{\gamma\delta})\phi & (\text{in } \mathcal{D}). \end{array}$$

In the last step the relation $(\widehat{\beta}\,\overline{\gamma\delta_1\ldots\delta_{d-1}})\phi(\delta_d\delta_{d+1})\phi = (\widehat{\beta}\,\overline{\gamma\delta})\phi$ is in $\mathcal{D}$ since

$$\begin{array}{rcl} |(\widehat{\beta}\,\overline{\gamma\delta_1\ldots\delta_{d-1}})\phi(\delta_d\delta_{d+1})\phi(\widehat{\beta}\,\overline{\gamma\delta})\phi| & = & |(\widehat{\beta}\,\overline{\gamma\delta_1\ldots\delta_{d-1}})\phi| + 1 + |(\widehat{\beta}\,\overline{\gamma\delta})\phi| \\ & \leq & |\widehat{\beta}| + |\overline{\gamma\delta_1\ldots\delta_{d-1}}| + 1 + |\widehat{\beta}| + |\overline{\gamma\delta}| \\ & \leq & 2M_\Sigma + 2M_\mathcal{B} + 1. \end{array}$$

Case 2.2: $d \in \{0,1\}$. This subcase splits into two further subcases depending on the value of c.

Case 2.2.1: $c \geq 2$. Since $\gamma \in \mathcal{PI}(A,T)$ no strict prefix of γ belongs to $\mathcal{L}(A,T)$. In particular we have $\gamma_1\ldots\gamma_{c-1} \in \mathcal{L}(A,U^1)$. Now we have:

$$\begin{array}{rcll} (\beta\gamma\delta)\phi & \equiv & (\beta\gamma_1\ldots\gamma_{c-1}\gamma_c\gamma_{c+1}\delta)\phi & \\ & \equiv & (\beta\gamma_1\ldots\gamma_{c-1})\phi(\gamma_c\gamma_{c+1}\delta)\phi & \text{(by Lemma 7.17 and since } c \geq 2) \\ & = & (\overline{\beta\gamma_1\ldots\gamma_{c-1}})\phi(\gamma_c\gamma_{c+1}\delta)\phi & \text{(by Lemma 7.19)} \\ & = & (\widehat{\beta}\,\overline{\gamma\delta})\phi & (\text{in } \mathcal{D}). \end{array}$$

In the last step the relation $(\overline{\beta\gamma_1\ldots\gamma_{c-1}})\phi(\gamma_c\gamma_{c+1}\delta)\phi = (\widehat{\beta}\,\overline{\gamma\delta})\phi$ belongs to $\mathcal{D}$ since

$$\begin{array}{rcl} |(\overline{\beta\gamma_1\ldots\gamma_{c-1}})\phi(\gamma_c\gamma_{c+1}\delta)\phi(\widehat{\beta}\,\overline{\gamma\delta})\phi| & = & |(\overline{\beta\gamma_1\ldots\gamma_{c-1}})\phi| + |(\gamma_c\gamma_{c+1}\delta)\phi| + |(\widehat{\beta}\,\overline{\gamma\delta})\phi| \\ & \leq & |\overline{\beta\gamma_1\ldots\gamma_{c-1}}| + 2 + |\widehat{\beta}| + |\overline{\gamma\delta}| \\ & \leq & M_{UU} + 2 + M_\Sigma + M_\mathcal{B}. \end{array}$$

Case 2.2.2: $c \in \{0,1\}$. In this case $c \in \{0,1\}$ and $d \in \{0,1\}$ and β has no prefix in $\mathcal{L}(A,T)$. It follows that $|(\beta\gamma\delta)\phi| \leq 2$ and so $(\beta\gamma\delta)\phi = (\widehat{\beta}\,\overline{\gamma\delta})\phi$ is in $\mathcal{D}$ since

$$|(\beta\gamma\delta)\phi(\widehat{\beta}\,\overline{\gamma\delta})\phi| \leq |(\beta\gamma\delta)\phi| + |(\widehat{\beta}\,\overline{\gamma\delta})\phi| \leq 2 + |\widehat{\beta}| + |\overline{\gamma\delta}| = 2 + M_\Sigma + M_\mathcal{B}$$

as required. □

Lemma 7.21. *The relations* $(\alpha\beta\gamma\delta)\phi = (\alpha\beta)\phi(\gamma\delta)\phi$ *where* $\alpha, \gamma \in \mathcal{PI}(A,T)$, $\beta, \delta \in \mathcal{L}(A,U^1), \alpha\beta \in \mathcal{L}(A,T)$ *and* $\gamma\delta \in \mathcal{L}(A,T)$ *are consequences of* $\mathcal{D}$.

Proof. First note that if β is the empty word then, by Lemma 7.17, we have:

$$(\alpha\beta\gamma\delta)\phi \equiv (\alpha\gamma\delta)\phi \equiv \alpha\phi(\gamma\delta)\phi \equiv (\alpha\beta)\phi(\gamma\delta)\phi.$$

Now suppose that β is non-empty. There are two cases to consider depending on whether or not $\beta\gamma\delta \in \mathcal{L}(A,T)$.

Case 1: $\beta\gamma\delta \notin \mathcal{L}(A,T)$. In this case $(\alpha\beta\gamma\delta)\phi$ is a single letter and the relation $(\alpha\beta\gamma\delta)\phi = (\alpha\beta)\phi(\gamma\delta)\phi$ is in $\mathcal{D}$ since

$$|(\alpha\beta\gamma\delta)\phi(\alpha\beta)\phi(\gamma\delta)\phi| = |(\alpha\beta\gamma\delta)\phi| + |(\alpha\beta)\phi| + |(\gamma\delta)\phi| \leq 1 + 2M_{\mathcal{B}}.$$

Case 2: $\beta\gamma\delta \in \mathcal{L}(A,T)$. In this case we have:

$$\begin{aligned}(\alpha\beta\gamma\delta)\phi &\equiv \alpha\phi(\beta\gamma\delta)\phi \\ &= \alpha\phi(\widehat{\beta}\,\overline{\gamma\delta})\phi \quad \text{(by Lemma 7.20)} \\ &= (\alpha\beta)\phi(\gamma\delta)\phi. \quad (\text{in } \mathcal{D})\end{aligned}$$

In the last step the relation $\alpha\phi(\widehat{\beta}\,\overline{\gamma\delta})\phi = (\alpha\beta)\phi(\gamma\delta)\phi$ is in $\mathcal{D}$ since

$$\begin{aligned}|\alpha\phi(\widehat{\beta}\,\overline{\gamma\delta})\phi(\alpha\beta)\phi(\gamma\delta)\phi| &\leq |\alpha\phi| + |(\widehat{\beta}\,\overline{\gamma\delta})\phi| + |(\alpha\beta)\phi| + |(\gamma\delta)\phi| \\ &\leq 1 + |\widehat{\beta}| + |\overline{\gamma\delta}| + 1 + 1 \\ &\leq 3 + |\widehat{\beta}| + |\overline{\gamma\delta}| \leq 3 + M_{\Sigma} + M_{\mathcal{B}}\end{aligned}$$

as required. □

Completing the proof of Theorem 7.2

We now complete the proof by proving that the relations (7.5) and (7.6) are all consequences of the fixed finite set of relations $\mathcal{D}$.

Lemma 7.22. *The relations $(w_1w_2)\phi = (w_1)\phi(w_2)\phi$ where $w_1, w_2 \in \mathcal{L}(A,T)$ are consequences of $\mathcal{D}$.*

Proof. We proceed by induction on the length of the word w_1w_2. When $|w_1w_2| \leq 2$ the relation $(w_1w_2)\phi = (w_1)\phi(w_2)\phi$ is in $\mathcal{D}$ since:

$$|(w_1w_2)\phi(w_1)\phi(w_2)\phi| = |(w_1w_2)\phi| + |w_1\phi| + |w_2\phi| \leq 2 + 1 + 1 = 4.$$

Let $w_1, w_2 \in \mathcal{L}(A,T)$ and suppose that the result holds for all $w_1', w_2' \in \mathcal{L}(A,T)$

such that $|w'_1 w'_2| < |w_1 w_2|$. Now decompose the word w_2:

$$w_2 \equiv \beta_1 \dots \beta_b \beta_{b+1}$$

where the principal factors are $\beta_1, \dots, \beta_{b-1}, \beta_b\beta_{b+1}$. Now consider the following prefixes of the word $w_1 w_2$:

$$\xi_0 \equiv w_1, \quad \xi_i \equiv w_1 \beta_1 \dots \beta_i \quad (2 \le i \le b-1).$$

These words all belong to $\mathcal{L}(A,T)$ since they are products of elements of $\mathcal{L}(A,T)$. There are two cases to consider:

Case 1: $|\xi_k \phi| = 1$ *for all* $0 \le k \le b-1$. In this case we can repeatedly apply Lemma 7.21 to get:

$$\begin{aligned}
(w_1\phi)(w_2\phi) &\equiv (w_1)\phi(\beta_1\beta_2 \dots \beta_{b-1}\beta_b\beta_{b+1})\phi \\
&\equiv (w_1)\phi(\beta_1)\phi(\beta_2)\phi \dots (\beta_{b-1})\phi(\beta_b\beta_{b+1})\phi \\
&= (w_1\beta_1)\phi(\beta_2)\phi \dots (\beta_{b-1})\phi(\beta_b\beta_{b+1})\phi \\
&= (w_1\beta_1\beta_2)\phi \dots (\beta_{b-1})\phi(\beta_b\beta_{b+1})\phi \\
&= \dots \\
&= (w_1\beta_1\beta_2 \dots \beta_{b-1})\phi(\beta_b\beta_{b+1})\phi \\
&= (w_1\beta_1\beta_2 \dots \beta_{b-1}\beta_b\beta_{b+1})\phi \\
&\equiv (w_1 w_2)\phi.
\end{aligned}$$

Case 2: $|\xi_k \phi| > 1$ *for some* $0 \le k \le b-1$. Let k be the smallest number such that $|\xi_k \phi| > 1$. Decompose ξ_k into principal factors:

$$\xi_k \equiv \gamma_1 \dots \gamma_c \gamma_{c+1}$$

where, since $|\xi_k \phi| > 1$, we know that $c \ge 2$. Proceeding as in Case 1 we first obtain

$$(w_1\phi)(w_2\phi) = (w_1\beta_1\beta_2 \dots \beta_k)\phi(\beta_{k+1} \dots \beta_b\beta_{b+1})\phi.$$

This time we continue as follows:

$$\begin{aligned}
& (w_1\beta_1\beta_2\ldots\beta_k)\phi(\beta_{k+1}\ldots\beta_b\beta_{b+1})\phi \\
\equiv\ & (\gamma_1\ldots\gamma_c\gamma_{c+1})\phi(\beta_{k+1}\ldots\beta_b\beta_{b+1})\phi \\
\equiv\ & (\gamma_1)\phi(\gamma_2\ldots\gamma_{c+1})\phi(\beta_{k+1}\ldots\beta_b\beta_{b+1})\phi && \text{(by Lemma 7.17 and since } c \geq 2) \\
=\ & (\gamma_1)\phi(\gamma_2\ldots\gamma_{c+1}\beta_{k+1}\ldots\beta_b\beta_{b+1})\phi && \text{(induction)} \\
\equiv\ & (\gamma_1\gamma_2\ldots\gamma_{c+1}\beta_{k+1}\ldots\beta_b\beta_{b+1})\phi && \text{(by Lemma 7.17)} \\
\equiv\ & (w_1w_2)\phi
\end{aligned}$$

as required. □

Lemma 7.23. *The relations* $(w_3xw_4)\phi = (w_3yw_4)\phi$ *where* $w_3, w_4 \in A^*, (x = y) \in \mathfrak{R}$ *and* $w_3xw_4 \in \mathcal{L}(A,T)$ *are consequences of* $\mathcal{D}$.

Proof. We proceed by induction on the combined length of w_3xw_4 and w_3yw_4. When $|w_3xw_4w_3yw_4| = 2$ the words w_3 and w_4 are empty and the relation $(w_3xw_4)\phi = (w_3yw_4)\phi$ belongs to $\mathcal{D}$ since

$$|(w_3xw_4)\phi(w_3yw_4)\phi| = |(w_3xw_4)\phi| + |(w_3yw_4)\phi| \leq |w_3xw_4| + |w_3yw_4| \leq 2.$$

Let $w_3, w_4 \in A^*, (x = y) \in \mathfrak{R}$ and $w_3xw_4 \in \mathcal{L}(A,T)$ and suppose that the result holds for all w_3', w_4' and $(x' = y')$ satisfying the analogous conditions where $|w_3'x'w_4'w_3'y'w_4'| < |w_3xw_4w_3yw_4|$. There are two cases to consider depending on whether or not w_3 has a prefix that belongs to $\mathcal{L}(A,T)$.

Case 1: *w_3 has a prefix that belongs to $\mathcal{L}(A,T)$.* Then write $w_3 \equiv w_3'w_3''$ where w_3' is the shortest such prefix. This case now splits into two subcases depending on whether or not $w_3''xw_4 \in \mathcal{L}(A,T)$.

Case 1.1: $w_3''xw_4 \notin \mathcal{L}(A,T)$. Then we also have $w_3''yw_4 \notin \mathcal{L}(A,T)$ and so $(w_3xw_4)\phi$ and $(w_3yw_4)\phi$ are both single letters and the relation $(w_3xw_4)\phi = (w_3yw_4)\phi$ is trivial.

Case 1.2: $w_3''xw_4 \in \mathcal{L}(A,T)$. Then we have:

$$\begin{aligned}
(w_3xw_4)\phi &\equiv (w_3'w_3''xw_4)\phi \\
&\equiv (w_3')\phi(w_3''xw_4)\phi && \text{(by Lemma 7.17)} \\
&= (w_3')\phi(w_3''yw_4)\phi && \text{(induction)} \\
&\equiv (w_3yw_4)\phi.
\end{aligned}$$

Case 2: *w_3 has no prefix that belongs to $\mathcal{L}(A,T)$.* In this case we decompose our words into principal factors:

$$w_3xw_4 \equiv w_3\beta_1\ldots\beta_{b+1}\gamma_1\ldots\gamma_{c+1}, \quad w_3yw_4 \equiv w_3\beta'_1\ldots\beta'_{b'+1}\gamma'_1\ldots\gamma'_{c'+1}$$

where $x \equiv \beta_1\ldots\beta_{b+1}$, $y \equiv \beta'_1\ldots\beta'_{b'+1}$ and the principal factors of w_3xw_4 are $w_3\beta_1, \beta_2, \ldots, \beta_b, \beta_{b+1}\gamma_1, \gamma_2, \ldots, \gamma_{c-1}$ and $\gamma_c\gamma_{c+1}$, and those of w_3yw_4 are $w_3\beta'_1$, $\beta'_2, \ldots, \beta'_{b'}$, $\beta'_{b'+1}\gamma'_1$, $\gamma'_2, \ldots, \gamma'_{c'-1}$ and $\gamma'_{c'}\gamma'_{c'+1}$. There are two cases to consider depending on the values of c and of c'.

Case 2.1: *$c \geq 2$ or $c' \geq 2$.* If $c \geq 2$ we have:

$$\begin{aligned}(w_3xw_4)\phi &\equiv (w_3x\gamma_1\ldots\gamma_c\gamma_{c+1})\phi \\ &\equiv (w_3x\gamma_1\ldots\gamma_{c-1})\phi(\gamma_c\gamma_{c+1})\phi \quad \text{(by Lemma 7.17)} \\ &= (w_3y\gamma_1\ldots\gamma_{c-1})\phi(\gamma_c\gamma_{c+1})\phi \quad \text{(induction)} \\ &= (w_3y\gamma_1\ldots\gamma_{c-1}\gamma_c\gamma_{c+1})\phi \quad \text{(by Lemma 7.22)} \\ &\equiv (w_3yw_4)\phi.\end{aligned}$$

The case $c' \geq 2$ is dealt with analogously.

Case 2.2: *$c, c' \in \{0,1\}$.* In this case first note that $|(w_3xw_4)\phi| = b + c \leq M_{\mathfrak{R}} + 1$. Likewise $|(w_3yw_4)\phi| = b' + c' \leq M_{\mathfrak{R}} + 1$ and we conclude that $|(w_3xw_4)\phi(w_3yw_4)\phi| \leq 2M_{\mathfrak{R}}+2$ and therefore the relation $(w_3xw_4)\phi = (w_3yw_4)\phi$ belongs to $\mathcal{D}$. □

7.5 Applications

In this section we list a number of corollaries of the main theorems of the chapter. Recall that, given a semigroup S and a subsemigroup T of S the Rees index of T in S is the cardinality of the set $S \setminus T$. If $S \setminus T$ is finite we say that T has finite Rees index in S.

Corollary 7.24. *If S is finitely generated (resp. presented) and T has finite Rees index in S then T is finitely generated (resp. presented).*

Proof. By Proposition 7.5, T has a finite boundary in S and is therefore by Theorems 7.1 and B finitely generated (resp. presented). □

This is in fact the more difficult direction of the main result of [82]. We remark here that the proof of the finite presentability result given in [82] is incomplete.

We outline briefly here the nature of the problem and explain how to fix it directly, without reference to Theorem 7.2.

Stage 2 of the proof of [82, Theorem 1.3] is concerned with proving that the relation $(\alpha\beta\gamma)\phi = (\alpha\overline{\beta}\gamma)\phi$ is a consequence of a suitably defined set of "short" relations $\mathcal{D}$, for all words α, β, γ such that $\alpha, \beta, \gamma \in A^*$ and $\beta \notin \mathcal{L}(A, T)$. The words $\alpha\beta\gamma$ and $\alpha\overline{\beta}\gamma$ are decomposed into principal factors as:

$$\alpha\beta\gamma \equiv \alpha_1 a_{i_1} \ldots \alpha_p a_{i_p} \alpha_{p+1} \beta_1 a_{j_1} \ldots \beta_q a_{j_q} \beta_{q+1} \gamma_1 a_{k_1} \ldots \gamma_r a_{k_r} \gamma_{r+1},$$

$$\alpha\overline{\beta}\gamma \equiv \alpha_1 a_{i_1} \ldots \alpha_p a_{i_p} \alpha_{p+1} \beta'_1 a_{l_1} \ldots \beta'_s a_{l_s} \beta'_{s+1} \gamma'_1 a_{m_1} \ldots \gamma'_t a_{m_t} \gamma'_{t+1}.$$

Here the notation is slightly different than in the previous section: each a-symbol, except for a_{k_r} and a_{m_t}, is the last symbol in its principal factor. Then six cases are considered depending on the values of the numbers p, q, r, s and t. The problems occur when either $r = 0$ or $t = 0$. For instance Case 6 deals with the situation when $p = r = t = 0$. In this case the decompositions become:

$$\alpha\beta\gamma \equiv \alpha_1 \beta_1 a_{j_1} \ldots \beta_q a_{j_q} \beta_{q+1} \gamma_1, \quad \alpha\overline{\beta}\gamma \equiv \alpha_1 \beta'_1 a_{l_1} \ldots \beta'_s a_{l_s} \beta'_{s+1} \gamma'_1$$

where the last principal factor in the first decomposition is $\beta_q a_{j_q} \beta_{q+1} \gamma_1$ and in the second decomposition is $\beta'_s a_{l_s} \beta'_{s+1} \gamma'_1$. The second step in the proof of Case 6 claims that the relation

$$(\pi(\epsilon, \alpha)\overline{\delta_1} a \overline{\delta_2} \overline{\gamma_1})\phi = (\pi(\epsilon, \alpha)\overline{\beta}\, \overline{\gamma'_1})\phi$$

belongs to $\mathcal{D}$. This is only true if both γ_1 and γ'_1 belong to $\mathcal{L}(A, U^1)$ otherwise we we have very little control over $|\overline{\gamma_1}|$ and $|\overline{\gamma'_1}|$. In order to deal with the cases where $\gamma_1 \in \mathcal{L}(A, T)$ or $\gamma'_1 \in \mathcal{L}(A, T)$ a new concept must be introduced.

Let $\alpha, \beta \in A^*$ where $\alpha \in \mathcal{L}(A, U^1)$ and $\alpha\beta \in \mathcal{L}(A, U^1)$. Amongst all the words β_1 with the property that $\alpha\beta_1 = \alpha\beta$ in S let $\sigma(\alpha, \beta)$ be one of minimal possible length. Now define the number

$$P^* = \max\{|\sigma(\alpha, \beta)| : \alpha, \alpha\beta \in \mathcal{L}(A, U^1)\}.$$

Since T has finite Rees index in S this number is well defined. It gives a measure of how far apart two elements of U can be in the right Cayley graph.

The definition of $\mathcal{D}$ must now be altered so that $\mathcal{D}$ also contains all relations that hold in T and whose length does not exceed

$$Q^* = \max\{5M + 4P + 3 + 2P^*, 7M + P + 4, 4M + N\}.$$

Now Case 6 splits into four parts depending on whether or not γ_1 or γ_1' belong to $\mathcal{L}(A,T)$. Changing the proof, in each case, involves replacing occurrences of $\overline{\gamma_1}$, when $\gamma_1 \in \mathcal{L}(A,T)$, in the original proof by $\sigma(\beta_{q+1}, \gamma_1)$. Similarly, whenever $\gamma_1' \in \mathcal{L}(A,T)$ we replace $\overline{\gamma_1'}$ by $\sigma(\beta'_{s+1}, \gamma_1')$. So for example if $\gamma_1 \in \mathcal{L}(A,T)$ and $\gamma_1' \notin \mathcal{L}(A,T)$ then the argument becomes:

$$\begin{aligned}(\alpha\beta\gamma)\phi &\equiv (\pi(\epsilon,\alpha)\beta_1 a_{j_1}\dots\beta_q a_{j_q}\beta_{q+1}\gamma_1)\phi \\ &\equiv (\pi(\epsilon,\alpha)\beta_1 a_{j_1}\dots\beta_q a_{j_q}\beta_{q+1}\sigma(\beta_{q+1},\gamma_1))\phi \\ &\equiv (\pi(\epsilon,\alpha)\delta_1 a\delta_2\sigma(\beta_{q+1},\gamma_1))\phi \\ &= (\pi(\epsilon,\alpha)\overline{\delta_1}a\overline{\delta_2}\sigma(\beta_{q+1},\gamma_1))\phi \\ &= (\pi(\epsilon,\alpha)\overline{\beta}\ \overline{\gamma_1'}) \equiv (\alpha\overline{\beta}\gamma)\phi.\end{aligned}$$

Using exactly the same method one may also fix the proofs of Cases 4 and 5.

Our second corollary concerns subsemigroups of free semigroups.

Corollary 7.25. *[18, Corollary 3.6] If F is a free semigroup with finite rank and I is a two-sided ideal that is finitely generated as a subsemigroup then I is finitely presented.*

Proof. We will show that I has a finite boundary in F. Let W be a finite generating set for I. If w belongs to the right boundary of I then, writing $w \equiv w'a$ where a is the last letter of w, the word w' belongs to $F \setminus I$. Since I is a right ideal no prefix of w belongs to I and therefore $w \in W$, a finite set. A similar argument tells us that the left boundary is finite. Then I is finitely presented by Theorem 7.2. □

In fact, in [18] a stronger result is proved: I must have finite Rees index in F.

In general, it is not true that ideals of finitely presented semigroups that are finitely generated as subsemigroups are finitely presented. One such example is given in [19, Theorem 3.1]. This example is in fact a semilattice of semigroups where the semilattice in question is a two element chain. In other words, it has the form $S = T \cup I$ where I is an ideal. We can obtain a positive result for this case if we introduce an additional hypothesis. If $S = T \cup I$ (a disjoint union) where T is a semigroup and I is an ideal then T acts naturally by left and right multiplication on I. The orbits of these actions are the sets iT^1 and $T^1 i$ where $i \in I$.

Corollary 7.26. *Let $S = T \cup I$, a disjoint union, where I is a two-sided ideal of S and T is a subsemigroup of S. If S is finitely generated (resp. presented) and the cardinality of every orbit of I under the action of T is finite then I is finitely generated (resp. presented).*

Proof. Since all the the orbits are finite it follows that properties (i) and (ii) of Proposition 7.6 both hold. Also, since T is a subsemigroup of S clearly $T^2 \cap I$ is finite (in fact it is empty) so condition (iii) of Proposition 7.6 holds. It now follows from Proposition 7.6 that T has a finite boundary in S and is therefore finitely generated (resp. presented) by Theorem 7.1 (resp. Theorem 7.2). □

We also have as a corollary the following folk-lore result concerning subsemigroups with ideal complement.

Corollary 7.27. *Let S be a semigroup with T a subsemigroup of S such that $S \setminus T$ is an ideal. If S is finitely generated (resp. presented) then T is finitely generated (resp. presented).*

Proof. For any finite generating set A of S the sets $UA \cap T$ and $AU \cap T$ are both empty since U is an ideal. It follows that the boundary of T in S has no more than $|A|$ elements and so is finite. The result then follows from Theorem 7.1 (resp. Theorem 7.2). □

The next example shows that it is possible to have a chain $K \leq T \leq S$ of subsemigroups such that the boundaries of K in T and of T in S are finite while the boundary of K in S is infinite. It therefore warns us against taking too far the analogy between boundaries and other notions of index.

Example 7.28. Let S be the semigroup with underlying set

$$S = [\mathbb{N}^0 \times \mathbb{N}^0 \times \mathbb{N}^0 \setminus \{(0,0,0)\}] \cup \{0\}$$

and multiplication that we describe below.

Define $F : S \setminus \{0\} \to \{1,2,3\}$ where, for $\alpha \in S \setminus \{0\}$, $F(\alpha)$ is the position of the first non-zero entry of the triple α (e.g. $F(0,1,1) = 2$). Also define $L : S \setminus \{0\} \to \{1,2,3\}$ where $L(\alpha)$ is the position of the last non-zero entry of α (e.g. $L(0,1,1) = 3$). Now multiplication in S is given by:

$$(x_1, y_1, z_1)(x_2, y_2, z_2) = \begin{cases} (x_1 + x_2, y_1 + y_2, z_1 + z_2) & \\ \qquad\qquad\qquad \text{if } L(x_1, y_1, z_1) \leq F(x_2, y_2, z_2) & \\ 0 \qquad\qquad\qquad\qquad \text{otherwise} & \end{cases}$$

and

$$(x, y, z)0 = 0(x, y, z) = 0 \cdot 0 = 0.$$

So for example $(1,2,3)(4,5,6) = 0$ since $L(1,2,3) = 3 > 1 = F(4,5,6)$. On the other hand $(1,2,0)(0,5,6)=(1,7,6)$ since $L(1,2,0)=2\leq 2=F(0,5,6)$. It is routine to check that the multiplication is associative.

The semigroup S is finitely generated by the set $A=\{(1,0,0),(0,1,0),(0,0,1)\}$ since for $(x,y,z)\in S\setminus\{0\}$ we can write

$$(x,y,z)=(1,0,0)^x(0,1,0)^y(0,0,1)^z$$

and we generate 0 with $(0,0,1)(1,0,0)=0$. In fact, S is defined by the presentation

$$\langle a,b,c,0\mid ba=0,cb=0,ca=0,0^2=a0=0a=b0=0b=c0=0c=0\rangle$$

where a,b and c correspond to the generators $(1,0,0),(0,1,0)$ and $(0,0,1)$, respectively. Let $B=\{(1,0,0),(0,1,0),(0,1,1)\}$, $C=\{(1,0,0),(1,1,1)\}$, $T=\langle B\rangle$ and $K=\langle C\rangle$. Clearly $K\leq T\leq S$. We begin with a straightforward observation concerning the elements of these subsemigroups:

$$T=\{(x,y,0):x\geq 0,y\geq 0,\ x \text{ and } y \text{ not both zero}\}$$
$$\cup\{(x,y,1):x\geq 0,y\geq 1\}\cup\{0\}$$

and

$$K=\{(x,0,0):x\geq 1\}\cup\{(x,1,1):x\geq 1\}\cup\{0\}.$$

Claim. *The boundary of K in S is infinite.*

Proof. We will show that $\mathcal{B}(A,K)$ in S is infinite. For all $x\geq 1$ we have $(x,1,0)\notin K$. However $(x,1,0)(0,0,1)=(x,1,1)\in K$ where $(0,0,1)\in A$. Thus $\{(x,1,1):x\geq 1\}$ is an infinite subset of the boundary of K in S. □

Claim. *The boundary of K in T is finite.*

Proof. We will show that $\mathcal{B}(B,K)$ in T is finite. Let $Q_1=\{(x,y,0):y\geq 1\}$, $Q_2=\{(x,y,1):y\geq 2\}$, and note that $T\setminus K=Q_1\cup Q_2\cup\{(0,1,1)\}$. Next note that every $(x,y,z)\in K$ has $y\leq 1$ so that the intersection of each of the sets

$$Q_2B,\quad BQ_2,\quad \{(0,1,0),(0,1,1)\}Q_1,\quad Q_1\{(0,1,0),(0,1,1)\},$$
$$\{(0,1,1)\}\{(0,1,0),(0,1,1)\},\quad \{(0,1,0),(0,1,1)\}\{(0,1,0)\}$$

with K is either empty or equal to $\{0\}$. Also

$$Q_1\{(1,0,0)\}=\{0\}=\{(0,1,1)\}\{(1,0,0)\}$$

by the definition of multiplication. Hence the right boundary is equal to $\{(1,0,0),0\}$.

For the left boundary we have

$$\begin{aligned} B(T\setminus K)^1 \cap K &= (B \cup (1,0,0)Q_1 \cup (1,0,0)\{(0,1,1)\} \cup \{0\}) \cap K \\ &= (B \cup \{(x,y,0) : x,y \geq 1\} \cup \{(1,1,1)\} \cup \{0\}) \cap K \\ &= \{(1,0,0),(1,1,1),0\}. \end{aligned}$$

We conclude that the boundary of K in T is equal to $\{(1,0,0),(1,1,1),0\}$. □

Claim. *The semigroup T has a finite boundary in S.*

Proof. We will show that $\mathcal{B}(A,T)$ in S is finite. Let $P_1 = \{(x,0,1) : x \geq 0\}$, $P_2 = \{(x,y,z) : z \geq 2\}$ noting that $S \setminus T = P_1 \cup P_2$. Note that every $(x,y,z) \in T$ has $z \leq 1$ and so the intersections of each of the sets

$$(0,0,1)P_1, \quad P_1(0,0,1), \quad P_2A \cap T, \quad AP_2 \cap T$$

with T is either empty or is equal to $\{0\}$. In addition $P_1(1,0,0) = P_1(0,1,0) = \{0\}$ and so the right boundary is equal to $\{0,(1,0,0),(0,1,0)\}$. For the left boundary we have

$$\begin{aligned} A(S\setminus T)^1 \cap T &= (A \cup (1,0,0)P_1 \cup (0,1,0)P_1) \cap T \\ &= (A \cup \{(x,0,1) : x \geq 1\} \cup \{(0,1,1),0\}) \cap T \\ &= \{(0,1,1),(1,0,0),(0,1,0),0\}. \end{aligned}$$

We conclude that the boundary of T in S is equal to $\{(1,0,0),(0,1,0),(0,1,1),0\}$. □

7.6 One sided boundaries

In §7.3 we saw that subsemigroups of finitely presented semigroups with only the right (or left) boundary finite need not be finitely generated, never mind finitely presented. This still leaves us with the question of whether finitely generated subsemigroups of finitely presented semigroups, with only a finite right (or left) boundary are always finitely presented. We now answer this question in the negative.

Let M be a monoid and let θ be an endomorphism of M. The Bruck–Reilly extension of M with respect to θ is the semigroup of triples $\mathbb{N}^0 \times M \times \mathbb{N}^0$ with multiplication defined by:

$$(m, a, n)(p, b, q) = (m - n + t, (a\theta^{t-n})(b\theta^{t-p}), q - p + t)$$

where $t = \max(n, p)$. Bruck–Reilly extensions are an important class of infinite simple semigroup. (For more details on Bruck–Reilly extensions see [57, Chapter 5].)

Proposition 7.29. *Suppose that the Bruck–Reilly extension $S = BR(M, \theta)$ of a monoid M is finitely generated and consider the subsemigroup $T = \{(0, a, n) : a \in M, n \in \mathbb{N}^0\}$. Then the right boundary of T in S is finite, while the left boundary is infinite.*

Proof. Let $U = \mathrm{BR}(M, \theta) \setminus T$. Since

$$m - n + t = m - n + \max(n, p) \geq m - n + n = m$$

it follows that U is a right ideal in $\mathrm{BR}(M, \theta)$ and thus, by Proposition 7.5, the right boundary of T in $\mathrm{BR}(M, \theta)$ is finite.

Let X be a finite generating set for $\mathrm{BR}(M, \theta)$. Since, by Proposition 7.3, the finiteness or otherwise of the left boundary is independent of the choice of generating set, we may assume without loss of generality that $(0, 1_M, 0) \in X$. For $n \in \mathbb{N}$ we have:

$$(0, 1_M, 1)(1, s, n) = (0, s, n);$$

note that here $(1, s, n) \in U$ and $(0, s, n) \in T$. Therefore, the left boundary of T in S is infinite (and equal to the whole of T). □

Example 7.30. Let M be a non-finitely presented monoid which has a finitely presented Bruck–Reilly extension $S = \mathrm{BR}(M, \theta)$. One possible choice for M is the group defined by the presentation:

$$\langle a, b, c, d \mid a^{2^i} b^{2^i} = c^{2^i} d^{2^i} \ (i \in \mathbb{N}^0) \rangle$$

where $\theta : M \to M$ extends the map $x\theta = x^2$ for $x \in \{a, b, c, d\}$. This example is taken from [83, Proposition 3.3] where it was shown that M is finitely generated but not finitely presented and $\mathrm{BR}(M, \theta)$ is finitely presented.

Let T be as in the proposition and let $N = \{(0, a, 0) : a \in M\}$. Clearly $N \cong M$, $N \subseteq T$ and $T \setminus N$ is an ideal of T. Hence, by Corollary 7.27, T is not finitely presented, although, it is finitely generated: any finite generating set for

N together with the element $(0, 1_M, 1)$ is a generating set for T. It now follows from Proposition 7.29 that T has a finite right boundary.

7.7 Left and right independence

We have seen three situations where the right and left boundaries are both finite. We also know that the right boundary may be infinite while the left finite and vice-versa, consider the boundaries in the bicyclic monoid for example (see Figure 7.2). The next proposition shows how any possible combination of left and right boundary sizes is possible.

Proposition 7.31. *Let S be the semigroup with set of generators $A = \{a\} \cup B \cup C \cup \{0\}$, where B and C are finite alphabets, and relations R given by:*

$$\begin{array}{lll} ab = b, \quad ba = 0 & & b \in B \\ ac = 0, \quad ca = c & & c \in C \\ x0 = 0x = 0 & & x \in A. \end{array}$$

Let $T = \langle A \setminus a \rangle$. Then:

(i) $|\mathcal{B}_r(A, T)| = |B| + 1$;

(ii) $|\mathcal{B}_l(A, T)| = |C| + 1$;

(iii) S and T are both infinite;

(iv) T has infinite Rees index in S.

Proof. First we find a normal form for the elements of S. We claim that every element of S may be written uniquely as a word from the set

$$N = \{0\} \cup \{a^i : i \in \mathbb{N}\} \cup B \cup C \cup BC \cup CB \cup BCB \cup CBC \cup \ldots \subseteq A^+.$$

If $w \in \{a^i : i \in \mathbb{N}\}$ then none of the relations of R may be applied to w and $w = u$ in S only if $u \equiv w$. If $w \equiv b_1c_1b_2c_2 \ldots b_pc_p$ then applying relations from R to w must give a word of the form

$$a^{i_1}b_1c_1a^{i_2}b_2c_2a^{i_3} \ldots a^{i_p}b_pc_pa^{i_{p+1}}$$

where $i_j \in \mathbb{N}^0$ for all j. If $w \equiv c_0b_1c_1b_2c_2 \ldots b_pc_p$ then applying relations from R to w must give a word of the form

$$c_0a^{i_1}b_1c_1a^{i_2}b_2c_2a^{i_3} \ldots a^{i_p}b_pc_pa^{i_{p+1}}$$

where $i_j \in \mathbb{N}^0$ for all j. If $w = b_1c_1b_2c_2 \dots b_pc_pb_{p+1}$ then applying relations from R to w must give a word of the form

$$a^{i_1}b_1c_1a^{i_2}b_2c_2a^{i_3} \dots a^{i_p}b_pc_pa^{i_{p+1}}b_{p+1}$$

where $i_j \in \mathbb{N}^0$ for all j. Also, If $w = c_1b_1c_2b_2 \dots c_pb_p$ then applying relations from R to w must give a word of the form

$$c_1a^{i_1}b_1c_2a^{i_2}b_2c_3a^{i_3} \dots a^{i_{p-1}}b_{p-1}c_pa^{i_p}b_p$$

where $i_j \in \mathbb{N}^0$ for all j. In particular if w has any of these four forms then $w \neq 0$ and $w \notin \{a^i : i \in \mathbb{N}\}$. It follows from the observations above that if $w, u \in N$ with $w \not\equiv u$ then $w \neq u$ in S, i.e. the words in N represent distinct elements.

Let $w \in A^+$. We have to show that using the relations from R the word w can be transformed into one of the words from the set N. If $w \in N$ then stop. Otherwise, if w contains the symbol 0, or has a subword from either of BA or AC then applying relations we can transform w into 0. Also, for all $b_1, b_2 \in B$ we have

$$b_1b_2 = b_1ab_2 = 0b_2 = 0.$$

Similarly for all $c_1, c_2 \in C$ we have $c_1c_2 = 0$. Therefore, if w has a subword from either the set BB or the set CC then applying relations we can transform w into 0. Otherwise w must belong to one of the sets:

$$\begin{aligned}
&a^*BCa^*BCa^* \dots a^*BCa^* \\
&Ca^*BCa^*BCa^* \dots a^*BCa^* \\
&a^*BCa^*BCa^* \dots a^*BCa^*B \\
&Ca^*BCa^*BCa^* \dots a^*BCa^*B
\end{aligned}$$

and any powers of a may be removed from w by applying relations from R, thus reducing w to a word from N.

Now that the normal forms have been established the rest of the proof is straightforward. The semigroup S is infinite since the set N is infinite. The subsemigroup T has normal forms

$$M = \{0\} \cup B \cup C \cup BC \cup CB \cup BCB \cup CBC \cup \dots \subseteq A^+$$

and is infinite since M is infinite. The Rees index of T in S is infinite since $N \setminus M$

is infinite. Computing the boundaries we have:

$$\mathcal{B}_r(A,T) = \{a^i : i \in \mathbb{N}^0\}A \cap T = B \cup \{0\},$$

$$\mathcal{B}_l(A,T) = A\{a^i : i \in \mathbb{N}^0\} \cap T = C \cup \{0\}.$$

Therefore $|\mathcal{B}_r(A,T)| = |B| + 1$ and $|\mathcal{B}_l(A,T)| = |C| + 1$. □

7.8 The converse: unions of semigroups

When defining the boundary $\mathcal{B}(A,T)$ it is essential to assume that S is finitely generated. Therefore the converse of Theorem 7.1 is not a sensible thing to consider. The converse of Theorem 7.2 may be formulated as follows. Let S be a semigroup generated by a finite set A and let T be a subsemigroup of S. If T is finitely presented and $\mathcal{B}(A,T)$ is finite then is S necessarily finitely presented? It is not hard to see that the answer to this question is no in general. For example, if S is any non-finitely presented semigroup that has a finite subsemigroup T then T is finitely presented and has a finite boundary in S.

One interesting situation where the converse does hold is when the complement of T happens to be a subsemigroup of S, i.e. when S is a disjoint union of two subsemigroups. In general we can prove the following result when S is a disjoint union of finitely many subsemigroups.

Theorem 7.32. *Let $S = \bigcup_{i\in I} S_i$, a disjoint union, where I is finite and each S_i is a subsemigroup of S. If each S_i is finitely presented and has a finite right boundary in S then S itself is finitely presented.*

Proof. For each $i \in I$ let S_i be defined by the presentation $\langle A_i | R_i\rangle$. We will write a presentation for S of the form $\langle B | \bigcup_{i\in I} R_i, R\rangle$ where $B = \bigcup_{i\in I} A_i$ and R is a finite set of relations holding in S that we describe below.

Let $i, j \in I$ with $i \neq j$. Consider the set of words $w \in A_i^+$ such that there exists some $a \in B$ with $wa \in \mathcal{L}(B, S_j)$. Denote this set of words by $W_r(i,j) \subseteq A_i^+$. Note that the elements that the words $W_r(i,j)$ represent may constitute an infinite subset of S_i. Let $w \in W_r(i,j)$ and $a \in B$ with $wa \in \mathcal{L}(B, S_j)$. Amongst all the words $w_1 \in A_i^+$ with the property that $wa = w_1 a$ in S let $\pi_r(w,a)$ be such a word of shortest length. So we have $wa = \pi_r(w,a)a$ in S. Now define:

$$d_r(i,j) = \max\{|\pi_r(w,a)| : w \in A_i^+, a \in B, wa \in \mathcal{L}(B, S_j)\}$$

provided $W_r(i,j)$ is non-empty; when $W_r(i,j)$ is empty we define $d_r(i,j) = 0$.

The number $d_r(i,j)$ is well defined since B is finite and the boundary of S_j in S is finite. Now define

$$f = \max\{d_r(i,j) : i, j \in I, i \neq j\}$$

which is well defined since I is finite. For every word $w \in B^+$ let $\widetilde{w} \in A_1^+ \cup \ldots \cup A_{|I|}^+$ be a fixed word such that $w = \widetilde{w}$ holds in S. Now let:

$$R = \{(w = \widetilde{w}) : w \in B^+, |w| \leq f + 2\}.$$

Note that the relations $wa = \pi_r(w,a)a$ with $|w| \leq f + 1$ are consequences of $\bigcup_{i \in I} R_i \cup R$.

Claim 1. *For every $w \in A_i^+$ and every $a \in B$ there exists some $u \in A_1^+ \cup \ldots \cup A_{|I|}^+$ such that $wa = u$ is a consequence of the relations R.*

Proof. We prove the claim by induction on the length of the word w. If $|w| \leq f+1$ then the relation $wa = \widetilde{wa}$ belongs to R and we are done. Now let $w \in A_i^+$ with $|w| > f + 1$ and suppose that the result holds for all $v \in A_i^+$ such that $|v| < |w|$. Write $wa \equiv w'w''a$ where $|w''| = f + 1$. There are two cases to consider:

Case 1: $w''a \in S_i$. Then the relation $w''a = \widetilde{w''a}$ belongs to R. Now $w' \in A_i^+$ and $\widetilde{w''a} \in A_i^+$ and so we can deduce

$$wa = w'(w''a) = w'(\widetilde{w''a}) \in A_i^+$$

as required.

Case 2: $w''a \in S_j$ *where* $j \neq i$. Then the relation $w''a = \pi_r(w'', a)a$ is a consequence of $\bigcup_{i \in I} R_i \cup R$ where $|\pi_r(w'', a)| = f < f + 1 = |w''|$ and so $|w'\pi_r(w'', a)| < |w'w''|$. Therefore we may apply induction to deduce:

$$wa = w'w''a = w'\pi_r(w'', a)a = u \in A_1^+ \cup \ldots \cup A_{|I|}^+$$

as required. □

Claim 2. *For every $w \in B^+$ there exists $u \in A_1^+ \cup \ldots \cup A_{|I|}^+$ such that $w = u$ is a consequence of the relations R.*

Proof. We prove the claim by induction on the length of the word w. When $|w| \leq f + 2$ the relation $w = \widetilde{w}$ belongs to R and we are done. Now let $w \in B^+$ with $|w| > f + 2$ and suppose that the result holds for all $v \in B^+$ with $|v| < |w|$.

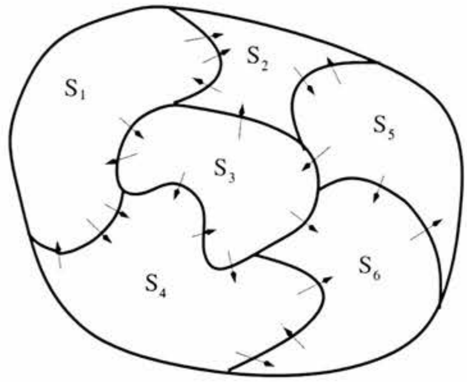


Figure 7.4: Picture of a semigroup that is the disjoint union of finitely many semigroups each with a finite boundary. It may be thought of as a patchwork quilt with finitely many patches sewn together with a finite amount of thread.

Write $w \equiv w'a$ where $a \in B$ is the last letter of w. By induction we can deduce $w' = u$ where $u \in A_1^+ \cup \ldots \cup A_{|I|}^+$. By Claim 1 we can deduce $ua = v$ where $v \in A_1^+ \cup \ldots \cup A_{|I|}^+$. It follows that we can deduce:

$$w \equiv w'a = ua = v \in A_1^+ \cup \ldots \cup A_{|I|}^+$$

as required. □

Let $w, v \in B^+$ such that $w = v$ holds in S and $w, v \in \mathcal{L}(B, S_i)$, say. Then there exist $w', v' \in A_i^+$ such that the relations $w = w'$ and $v = v'$ are consequences of R. Furthermore, the relation $w' = v'$ is a consequence of the relations R_i. Therefore, using the relations $\bigcup_{i \in I} R_i \cup R$ we may deduce $w = w' = v' = v$ and, since w and v were arbitrary, S is defined by the presentation $\langle B | \bigcup_{i \in I} R_i, R \rangle$. □

There is an obvious dual result where the left boundaries are all finite. Now if we combine Theorem 7.32 with Theorem 7.2 we obtain:

Corollary 7.33. *Let S be a finitely generated semigroup which can be decomposed into a finite disjoint union $S = \bigcup_{i \in I} S_i$ of subsemigroups with finite boundaries. Then S is finitely presented if and only if all the S_i are finitely presented.*

Example 7.34. Let $S_1, S_2, \ldots, S_m$ be finitely generated semigroups. Define a semigroup

$$T = S_1 \sqcup S_2 \sqcup \ldots \sqcup S_m \sqcup \{0\}$$

where all products of elements belonging to different S_i equal zero. We call T the 0-*disjoint union of* $S_1, \ldots, S_m$. Each of the semigroups S_i has a finite boundary in T, since they each have ideal complement. Also, the subsemigroup $\{0\}$ is finite and hence has a finite boundary. Therefore, T is a finitely generated semigroup which can be decomposed into a finite disjoint union of subsemigroups with finite boundaries. It now follows from Corollary 7.33 that T is finitely presented if and only if all the S_i are finitely presented.

Note that the converse of Theorem 7.32 does not hold in general. For example let $A = \{a, b\}$, $T_1 = \{aw : w \in A^*\}$ and $T_2 = \{bw : w \in A^*\}$. Then $S = A^+$ is the disjoint union of T_1 and T_2, both T_1 and T_2 have finite right boundaries in S but neither of them is finitely presented (since they are not even finitely generated).

Without the restriction that the boundaries are finite Corollary 7.33 no longer holds. For example, in [5, Example 3.4] an example is given of a non-finitely presented semigroup S that is a disjoint union of two finitely presented semigroups.

7.9 Subsemigroups of free semigroups

Earlier in this chapter we saw an application of Theorem 7.2 to finitely generated ideals of free semigroups with finite rank. We discuss a few other results relating to boundaries of subsemigroups of free semigroups in this section. Let $F = A^+$ where $|A| = r$ be the free semigroup of rank r.

Proposition 7.35. *If I is a right ideal in F and I is finitely generated as a subsemigroup then I has a finite right boundary in F.*

Proof. Let W be a finite generating set for I. Let w be on the right boundary of I in F with respect to the generating set A. This means that we can write $w \equiv ua$ where $a \in A$, $u \in A^+$ and u is not in T. Since I is a right ideal we conclude that no prefix of u belongs to T and therefore $w \in W$. □

We know that subsemigroups of free semigroups can have finite boundaries while having infinite Rees index. The next result tells us that the same is not true for the ideals of F.

Proposition 7.36. *A right (respectively left) ideal I of F has finite Rees index if and only if it has a finite left (respectively right) boundary in F.*

Proof. If the Rees index is finite then the boundaries are obviously finite. For the converse, suppose that I does not have finite index in F. Let Y be an infinite subset of $F \setminus I$ and let $w \in I$ be fixed. Then, since I is a right ideal, wY is an infinite subset of I. Clearly this is only possible if the left boundary of I in F is infinite. □

This gives the following corollary which was originally proven in [18].

Corollary 7.37. *If I is a two-sided ideal of F that is finitely generated as a subsemigroup then I has finite Rees index in F.*

Proof. Since I is a left ideal and is finitely generated, by Proposition 7.35, I has a finite left boundary in F. Therefore I is a right ideal with finite left boundary which, by Proposition 7.36, must have finite Rees index. □

There do exist examples of right ideals that are finitely generated but do not have finite index. It was shown in [18] that right ideals of F that are finitely generated as subsemigroups are finitely presented. The propositions above tell us that finite boundary theory alone cannot be used to deduce this result. That is, it is not true that every right ideal that is finitely generated as a subsemigroup has finite left and right boundaries.

There is an interesting question here though. We know that if T is a subsemigroup of F and T has a finite right boundary then T need not be finitely generated. For example given $a \in A$ the subsemigroup of all words $\{aw : w \in A^+\}$ is not finitely generated but has a finite right boundary. What happens if we force the subsemigroup to be finitely generated? We saw in Proposition 7.29 that, in general, finitely generated subsemigroups with finite one-sided boundaries will not be finitely presented. We will see below that for subsemigroups of free semigroups they are.

Given a word $w \in A^+$ define

$$\mathrm{pref}(w) = \{u : w \equiv uv \text{ for some } v \in A^*\}$$

and for a subset X of A^+ define

$$\mathrm{pref}(X) = \bigcup_{w \in X} \mathrm{pref}(w).$$

We call the set $\mathrm{pref}(X)$ the *prefix closure* of the set X. Given a subset X of A^+ we say that X is *prefix closed* if $\mathrm{pref}(X) = X$ and *almost prefix closed* if $\mathrm{pref}(X) \setminus X$ is finite. Also, given $w \in A^+$ and $n \in \{1, \ldots, |w|\}$. Then define $w(n) = a_1 \ldots a_n$ where $w \equiv a_1 \ldots a_{|w|}$.

The connection with subsemigroups with finite boundaries is given by the following result.

Proposition 7.38. *Let T be a subsemigroup of the free semigroup F_r of rank r. Then T has a finite right (resp. left) boundary if and only if T is almost prefix* *(resp. suffix) closed.*

Proof. Suppose that T is almost prefix closed. Let $w \in \mathcal{B}_r(A,T)$ where $|w| = m$. Then $w(m-1)$ belongs to $\mathrm{pref}(T) \setminus T$ which is finite, because T is almost prefix closed, and puts a bound on $|w|$. For the converse suppose that T is not almost prefix closed. Then we can find arbitrarily long words $w \in F \setminus T$ such that for some $u \in A^+$ we have $wu \in T$. This implies that we can find arbitrarily long words on the right boundary which, since F is free, contradicts the assumption that the right boundary is finite. □

Theorem 7.39. *Let F be a free semigroup of finite rank and let T be a finitely generated subsemigroup of F.*

(i) If T has a finite right boundary in F then T is finitely presented.

(ii) If T has a finite left boundary in F then T is finitely presented.

Proof. We prove part (i). Part (ii) follows from a dual argument.

By Proposition 7.38 there exists some $M \in \mathbb{N}$ such that for every $w \in T$ with $|w| \geq M$, and for every $l \geq M$, we have $w(l) \in T$. Fix a finite generating set $W = \{w_1, \ldots, w_m\}$ for T and let $B = \{b_1, \ldots, b_m\}$ be a new set of symbols in one-one correspondence with the words in W. Given a relation

$$b_{i_1} \ldots b_{i_r} = b_{j_1} \ldots b_{j_s}$$

that holds in T we define its length to be

$$|w_{i_1} \ldots w_{i_r}| = |w_{j_1} \ldots w_{j_s}|.$$

Define

$$N = \max\{|w| : w \in W\}.$$

We claim that T is defined by the presentation $\langle B \mid Q \rangle$ where $Q \subseteq B^+ \times B^+$ is the set of all relations that hold in T and have length not exceeding MN.

Let $(\alpha = \beta) \in B^+ \times B^+$ be a relation that holds in T. We prove the result by induction on the length of the relation $\alpha = \beta$. For the base case, if $(\alpha = \beta) \in Q$ then we are done. Now suppose that $(\alpha = \beta) \notin Q$ and that every relation that is strictly shorter can be deduced from the relations Q. Say

$$\alpha \equiv b_{i_1} \ldots b_{i_r}, \quad \beta \equiv b_{j_1} \ldots b_{j_s}.$$

Since $\alpha = \beta$ holds in T it follows that $w_{i_1} \dots w_{i_r} \equiv w_{j_1} \dots w_{j_s}$ in A^+. Suppose, without loss of generality, that $|w_{i_1}| \leq |w_{j_1}|$. There are two cases to consider depending on the length of the word w_{i_1}.

Case 1: $|w_{i_1}| = |w_{j_1}|$. It follows that $b_{i_1} \equiv b_{j_1}$ and $w_{i_2} \dots w_{i_r} \equiv w_{j_2} \dots w_{j_s}$ and so the relation $b_{i_2} \dots b_{i_r} \equiv b_{j_2} \dots b_{j_s}$ holds in T and has length strictly less than the relation $\alpha = \beta$. Therefore by induction we can deduce:

$$b_{i_1} b_{i_2} \dots b_{i_r} = b_{j_1} b_{i_2} \dots b_{i_r} = b_{j_1} b_{j_2} \dots b_{j_s}.$$

Case 2: $|w_{i_1}| < |w_{j_1}|$. Begin by writing $w_{j_1} \equiv w_{i_1}\zeta$ so that $w_{i_2} \dots w_{i_r} \equiv \zeta w_{j_2} \dots w_{j_s} \equiv w'$. Since $(\alpha = \beta) \notin Q$ it follows that $|w'| \geq M$ and $w' \in T$. By definition of M it follows that for every $l \geq M$ we have $w'(l) \in T$. It follows that for some $k \leq M$ we have $\zeta w_{j_2} \dots w_{j_k} \in T$. Therefore

$$|\zeta w_{j_2} \dots w_{j_k}| = |\zeta| + |w_{j_2}| + \ldots + |w_{j_k}| \leq kN \leq MN.$$

We can, therefore, write it in terms of generators giving:

$$\zeta w_{j_2} \dots w_{j_k} \equiv w_{l_1} \dots w_{l_q}.$$

The relation:

$$b_{j_1} b_{j_2} \dots b_{j_k} = b_{i_1} b_{l_1} \dots b_{l_q}$$

belongs to Q since it has length

$$|w_{j_1} \dots w_{j_k}| \leq kN \leq MN.$$

Moreover, the relation

$$b_{i_2} b_{i_3} \dots b_{i_r} = b_{l_1} \dots b_{l_q} b_{j_{k+1}} \dots b_{j_s}$$

holds in T and is deducible from the relations Q by induction since it has length

$$|w_{i_2} \dots w_{i_r}| < |w_{i_1} \dots w_{i_r}|.$$

Therefore we may deduce:

$$b_{j_1} b_{j_2} \dots b_{j_k} b_{j_{k+1}} \dots b_{j_s} = b_{i_1} b_{l_1} \dots b_{l_q} b_{j_{k+1}} \dots b_{j_s} = b_{i_1} b_{i_2} \dots b_{i_r}$$

as required. $\square$

As an immediate corollary we have the following result:

Corollary 7.40. *[18, Theorem 4.3] If R is a finitely generated right ideal of a free semigroup S, then R is finitely presented.*

So the finitely generated subsemigroups of F_r with finite one sided boundaries provide a fairly large class of finitely presented subsemigroups of F_r. By no means, however, does this set account for all finitely presented subsemigroups of F_r. For example $T = \langle ab \rangle \leq \{a, b\}^+$ is free, and hence finitely presented, but has an infinite right and infinite left boundary in T.

Open Problem 9. Use the notion of boundary to "describe" the subsemigroups of free semigroups that are finitely presented.

In [73] a finitely generated semigroup S isomorphic to some subsemigroup of a free semigroup is called an F-semigroup. In this paper Markov proves that it is decidable whether or not an F-semigroup has a finite presentation.

Chapter 8

Strict boundaries and unitary subsemigroups

8.1 Introduction

In the previous chapter the importance of the strict right and left boundaries became apparent. In particular, strict boundary elements played a crucial role when decomposing words from $\mathcal{L}(A,T)$ into principal factors. In this chapter we investigate the strict boundary in more detail.

The chapter is structured as follows. In §8.2 we recall the definitions of strict right and left boundaries given in the previous chapter and we provide some motivation for their study. In §8.3 we consider generating sets of unitary subsemigroups with finite strict boundaries and in §8.4 we consider presentations of such semigroups and present the main results of the chapter. Finally in §8.5 we give applications of the main results firstly to subgroups of groups and secondly to maximal subgroups of completely 0-simple semigroups.

8.2 Preliminaries

We first recall the definitions of strict right and left boundaries that were given in the previous chapter.

Definition 8.1. Let S be a semigroup generated by the finite set A and let T be a subsemigroup of S. 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 $\mathcal{L}(A,T)$. Also, 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, respectively, by $\mathcal{SWB}_r(A,T)$ and $\mathcal{SWB}_l(A,T)$.

Definition 8.2. The *strict left boundary* of T in S with respect to the generating set A is $\mathcal{SB}_l(A,T) = \mathcal{SWB}_l(A,T)\theta$ and the *strict right boundary* is defined as $\mathcal{SB}_r(A,T) = \mathcal{SWB}_r(A,T)\theta$. Moreover we define the *strict two-sided boundary* by $\mathcal{SB}(A,T) = \mathcal{SB}_l(A,T) \cup \mathcal{SB}_r(A,T)$.

The first thing to observe is that $\mathcal{SB}_l(A,T) \subseteq \mathcal{B}_l(A,T)$ and $\mathcal{SB}_r(A,T) \subseteq \mathcal{B}_r(A,T)$. This inclusion may well be proper as we will see below.

Unfortunately the finiteness or otherwise of the strict right, left and two-sided boundaries does depend on the choice of generating set.

Example 8.3. 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 $T =$

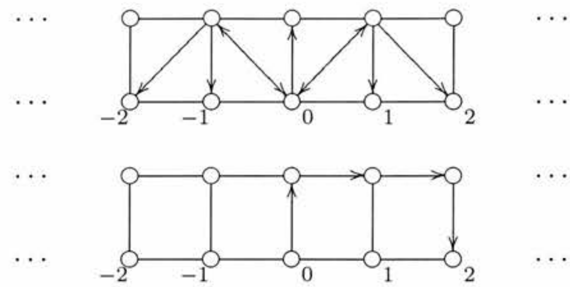


Figure 8.1: The right Cayley graph of S with respect to the generating sets A and B.

$\mathbb{Z} \oplus \{0\}$, a subsemigroup of S. 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 to $\{(0,0),(-2,0),(2,0),(1,0),(-1,0)\}$.

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

$$(0,1) + \underbrace{(1,0) + (1,0) + \ldots + (1,0)}_{m} + (0,1) = (m,0)$$

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

This leads to the following definition.

Definition 8.4. We say that T has a finite strict left, right or two-sided boundary in S if for some finite generating set A of S the sets $\mathcal{SB}_l(A,T)$, $\mathcal{SB}_r(A,T)$ or $\mathcal{SB}(A,T)$, respectively, are finite.

The motivation for the study of strict boundaries came from considering certain subgroups of groups. The main result of the previous chapter does not tell us anything interesting when applied to groups. Indeed, if G is an infinite group and H is a proper subgroup of G then, by Proposition 7.6(iii), H has a finite boundary in G if and only if H is finite. On the other hand, if G is a group and $N \trianglelefteq G$ such that the factor group G/N is a finite cyclic group then, as we will see in Lemma 8.21, N has a finite strict boundary in G.

Words in subgroups of groups decompose into principal factors in a particularly "nice" way. Let G be a group generated, as a semigroup, by the finite set A. Let H be a subgroup of G. Given a word $w \in A^+$ when we decompose it into principal factors (as in Lemma 7.17) we obtain:

$$w \equiv \alpha_1 \ldots \alpha_k \alpha_{k+1}$$

where $\alpha_i \in \mathcal{SB}_r(A,H)$ for $1 \leq i \leq k$ and $\alpha_{k+1} \equiv \epsilon$. This is because if $\alpha_{k+1} \in \mathcal{L}(A, G \setminus H)$ and $\alpha_k \in \mathcal{L}(A,H)$ then $\alpha_k\alpha_{k+1} \notin H$, contrary to assumption. It follows that the strict right boundary of a subgroup is a generating set for that subgroup. In the special cases when this strict boundary is finite we are able to conclude that the subgroup is finitely generated. More generally than this, the same is true for left and right unitary subsemigroups of arbitrary semigroups.

Definition 8.5. Let S be a semigroup with T a subsemigroup of 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

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

The main aim of this chapter is to prove the following result and discuss a couple of applications.

Theorem 8.6. *Let S be a semigroup and let T be a left unitary subsemigroup of S. If S is finitely generated (resp. presented) and T has a finite strict right boundary in S then T is finitely generated (resp. presented).*

There is an obvious dual result when T is right unitary and has a finite strict left boundary.

8.3 Generating sets

Our first result concerns generation.

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. By Proposition 7.7 the subsemigroup T is generated by $\mathcal{SB}_r(A,T)U^1 \cap T$. Since T is left unitary it follows that:

$$\mathcal{SB}_r(A,T)U^1 \cap T = \mathcal{SB}_r(A,T) \cap T = \mathcal{SB}_r(A,T).$$

Therefore, $\mathcal{SB}_r(A,T)$ is a finite generating set for T. □

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.

Notice that, unlike in the the previous section, the strict boundary does form a generating set for the subsemigroup. It is not true, in general, that subsemigroups with finite strict boundaries are finitely generated.

Example 8.9. Let $S = A^+$, where $A = \{a, b\}$, and $T = \{a\} \cup a\{a, b\}^*a$, a subsemigroup of S. Clearly T is not finitely generated since the elements ab^ia must all be included in any generating set. We have

$$\mathcal{SB}_l(A, T) = \mathcal{SB}_r(A, T) = \{a\}.$$

Therefore, T has a finite strict boundary in S but T is not finitely generated. Note that T is not left unitary since $a \in T$, $b^ia \notin T$ but $ab^ia \in T$. Similarly T is not right unitary.

8.4 Presentations

In this section we will prove the following result.

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 a finite strict right boundary in S then T is finitely presented;

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

We will go through an argument that is similar to the one given in the proof of Theorem 7.2 of the previous section. This time, however, the decomposition of words in $\mathcal{L}(A, T)$ is much less complicated and as a result the proof is more straightforward than that of Theorem 7.2. It is also worth noting that Theorem 8.10 neither implies, nor is implied by, Theorem 7.2 of the previous section.

Let S be a finitely presented semigroup with T a left unitary subsemigroup of S with a finite strict right boundary. Let A be a finite generating set for S with respect to which T has a finite strict right boundary, and let $\mathfrak{P} = \langle A | \mathfrak{R} \rangle$

be a finite presentation for S with respect to this generating set. Let η be the smallest congruence on A^+ containing $\mathfrak{R}$.

We will use the generating set described in Theorem 8.7. Define

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

a set of shortest length word representatives of the elements of the strict right boundary. Define a new alphabet in one to one correspondence with these generating words:

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

Let $\psi : B^+ \to A^+$ be the unique homomorphism extending $b_v \mapsto v$ and call this the representation mapping.

Now we define our rewriting mapping. For $w \in \mathcal{L}(A,T)$ write $w \equiv \alpha\beta$ where α is the shortest prefix of w that belongs to $\mathcal{L}(A,T)$. Since T is left unitary it follows that $\beta \in \mathcal{L}(A,T)$. We can, therefore, define ϕ inductively by

$$w\phi = \begin{cases} b_{\bar{\alpha}} \text{ if } \beta = \epsilon \\ b_{\bar{\alpha}}(\beta\phi) \text{ otherwise,} \end{cases}$$

where bar is defined as in Chapter 7, §7.4. Note that it is necessary to bar words in the above definition because it is possible that $\mathcal{SWB}_r(A,T)$ is infinite while $\mathcal{SB}_r(A,T)$ is finite.

Now by Theorem 7.14 it follows that the semigroup T is defined by the presentation with generators B and relations

$$b_v = v\phi \tag{8.1}$$

$$(w_1w_2)\phi = (w_1\phi)(w_2\phi) \tag{8.2}$$

$$(w_3xw_4)\phi = (w_3yw_4)\phi \tag{8.3}$$

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 $w_3xw_4 \in \mathcal{L}(A,T)$.

The relations (8.1) are all trivial by the definition of ϕ. We now prove that the relations (8.2) are also all trivial.

Lemma 8.11. *Let $w \in \mathcal{L}(A,T)$. Then w may be written uniquely as $w \equiv \alpha_1 \ldots \alpha_k$ where $k \geq 1$ and $\alpha_i \in \mathcal{L}(A, \mathcal{SB}_r(A,T))$ for all i. Moreover, this decom-*

position satisfies:

$$w\phi \equiv (\alpha_1 \ldots \alpha_k)\phi \equiv (\alpha_1\phi)\ldots(\alpha_k\phi) \equiv b_{\overline{\alpha_1}} \ldots b_{\overline{\alpha_k}}.$$

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

$$(w_1 w_2)\phi \equiv (\alpha_1 \ldots \alpha_k \beta_1 \ldots \beta_l)\phi \equiv b_{\overline{\alpha_1}} \ldots b_{\overline{\alpha_k}} b_{\overline{\beta_1}} \ldots b_{\overline{\beta_l}} \equiv (w_1\phi)(w_2\phi).$$

Proof. The first part follows directly from the definition of ϕ. For the second part let w_1 and w_2 be as in the statement of the lemma. It follows that $\alpha_i \in \mathcal{SB}_r(A,T)$ and $\beta_i \in \mathcal{SB}(A,T)$ for all i. Then from the definition of ϕ we have

$$(\alpha_1 \ldots \alpha_k \beta_1 \ldots \beta_l)\phi \equiv (\alpha_1\phi)\ldots(\alpha_k\phi)(\beta_1\phi)\ldots(\beta_l\phi)$$

and the result follows. □

We now write a presentation for T.

Theorem 8.12. *The presentation* $\langle B|Q\rangle$ *where*

$$Q = \{(u = v) \in B^+ \times B^+ : |uv| \leq \max\{|\alpha\beta| : (\alpha = \beta) \in \mathfrak{R}\}\}$$

defines T. *In particular, if* $\mathfrak{R}$ *is finite then* T *is finitely presented.*

As a consequence of the discussion above it is sufficient for us to prove the following lemma.

Lemma 8.13. *The relations* $(w\alpha v)\phi = (w\beta v)\phi$ *where* $w, v \in A^*$ *and* $(\alpha = \beta) \in \mathfrak{R}$ *are consequences of* Q.

Proof. We prove this by induction on the length of the word $w\alpha v w\beta v$. When $|w\alpha v| + |w\beta v| = 2$ we have $|(w\alpha v)\phi(w\beta v)\phi| = 2 = |\alpha\beta|$ and so, by definition, the relation $(w\alpha v)\phi = (w\beta v)\phi$ belongs to Q. Now suppose that the result holds for all $u', v' \in A^*$ and $(\alpha' = \beta') \in \mathfrak{R}$, satisfying the analogous conditions, such that $|w'\alpha' v' w'\beta' v'| < |w\alpha v w\beta v|$. Decompose $w\alpha v$ and $w\beta v$ into principal factors:

$$w\alpha v \equiv \gamma_1 \ldots \gamma_k, \quad w\beta v \equiv \delta_1 \ldots \delta_l.$$

Case 1: $|\gamma_1| \leq |w|$. Write $w \equiv \gamma_1\zeta$ where $\zeta \in A^*$. Then we have

$$\begin{aligned} (w\alpha v)\phi &\equiv (\gamma_1\zeta\alpha v)\phi \equiv \gamma_1\phi(\zeta\alpha v)\phi && \text{(Lemma 8.11)} \\ &= \gamma_1\phi(\zeta\beta v)\phi && \text{(by induction)} \\ &\equiv (\gamma_1\zeta\beta v)\phi \equiv (w\beta v)\phi. \end{aligned}$$

Case 2: $|\gamma_k| \leq |v|$. Write $v \equiv \zeta\gamma_k$ where $\zeta \in A^*$. Then we have

$$\begin{aligned} (w\alpha v)\phi &\equiv (w\alpha\zeta\gamma_k)\phi \equiv (w\alpha\zeta)\phi\gamma_k\phi && \text{(Lemma 8.11)} \\ &= (w\beta\zeta)\phi\gamma_k\phi && \text{(by induction)} \\ &\equiv (w\beta\zeta\gamma_k)\phi \equiv (w\beta v)\phi. \end{aligned}$$

Case 3: $|\delta_1| \leq |w|$. Exactly the same argument as in Case 1.

Case 4: $|\delta_l| \leq |v|$. Exactly the same argument as in Case 2.

Case 5: $|\gamma_1| > |w|$, $|\delta_1| > |w|$, $|\gamma_k| > |v|$ *and* $|\delta_l| > |v|$. It follows that $|\gamma_2 \ldots \gamma_{k-1}| \leq |\alpha| - 2$ and that $|\delta_2 \ldots \delta_{l-1}| \leq |\beta| - 2$. Therefore $k \leq |\alpha|$ and $l \leq |\beta|$. In particular $|(w\alpha v)\phi(w\beta v)\phi| = k + l \leq |\alpha\beta|$ and so $(w\alpha v)\phi = (w\beta v)\phi$ belongs to Q. □

In order to prove the dual result we decompose words into principal factors from right to left rather than from left to right. This completes the proof of Theorem 8.12.

The condition that the subsemigroup is left unitary cannot be removed. We saw in Example 8.9 that S can be finitely presented and have a subsemigroup T that has a finite strict boundary but is not finitely generated. Also, if S is finitely presented, $T \leq S$ is finitely generated and T has a finite strict boundary it is not necessarily true that T is finitely presented. To see this we consider subsemigroups of free semigroups.

Lemma 8.14. *Let $S = A^+$ where A is a finite alphabet and let T be a subsemigroup of S. If T is finitely generated then T has a finite strict boundary in S*

Proof. We will show that $\mathcal{SB}(A,T)$ is finite. Let $B \subseteq A^+$ be a finite generating set for T. Let $w \in \mathcal{SB}_r(A,T)$. Write $w \equiv \beta_1 \ldots \beta_k$ where $\beta_i \in B$ for all i. Since no strict prefix of w belongs to T it follows that $k = 1$ and $w \in B$. By a dual argument it follows that $\mathcal{SB}_l(A,T) \subseteq B$. Therefore $\mathcal{SB}(A,T) \subseteq B$ is finite. □

Example 8.15. Let T be a finitely generated and non-finitely presented subsemigroup of the free semigroup $S = A^+$, over a finite alphabet A (see Example 7.12 for example). Then T has a finite strict boundary in S by Lemma 8.14. On the other hand, since T is not finitely presented, it cannot have a finite boundary by Theorem 7.2.

Combining Lemma 8.14 with the presentation given in Theorem 8.12 we recover the following well known fact.

Corollary 8.16. *Finitely generated right or left unitary subsemigroups of free semigroups are free.*

In fact, a more general result than this is known. In Chapter 7 of [57], which is on the subject of codes, Proposition 7.2.1 states that a subsemigroup T of A^* is free if and only if

$$(\forall w \in A^+)[wT \cap T \neq \varnothing \ \&\ Tw \cap T \neq \varnothing] \Rightarrow w \in T.$$

In particular, left and right unitary subsemigroups satisfy this condition. This result is originally due to Schützenberger; see [86].

Of course, in general left and right unitary subsemigroups of finitely generated (presented) semigroups need not be finitely generated (presented). Consider subgroups of groups for example. We also have the following example.

Example 8.17. Recall the bicyclic monoid, introduced in Example 7.4. It is an inverse semigroup and so its idempotents form a subsemigroup. The subsemigroup of idempotents of this inverse semigroup is both left and right unitary. In fact this semigroup is an E-unitary inverse semigroup so this must be the case. This subsemigroup is isomorphic to an infinite semilattice (in fact an infinite chain) which is clearly not finitely generated. Both the left and right strict boundaries of E in B are infinite (see Figure 7.2).

8.5 Applications

Subgroups of groups

Since the original motivation for studying the strict boundary in more detail came from considering subgroups of groups we revisit them here. We begin with a result that compares the left, right and two-sided strict boundaries of a subgroup.

Proposition 8.18. *Let G be a group generated, as a semigroup, by a finite set A. Let H be a subgroup of G. Then:*

$$\mathcal{SB}_r(A,H) = \mathcal{SB}_l(A,H) = \mathcal{SB}(A,H).$$

Proof. Let $h \in \mathcal{SB}_r(A,H)$. Write $h = a_1 \ldots a_k \equiv w$ such that no strict prefix of w belongs to H. For all $r \in \{2, \ldots, k\}$ we have

$$a_r \ldots a_k = (a_1 \ldots a_{r-1})^{-1} h \notin H$$

since $a_1 \ldots a_{r-1} \notin H$ which implies that $(a_1 \ldots a_{r-1})^{-1} \notin H$. Therefore, $h \in \mathcal{SB}_l(A,H)$ and so $\mathcal{SB}_r(A,H) \subseteq \mathcal{SB}_l(A,H)$. Similarly $\mathcal{SB}_l(A,H) \subseteq \mathcal{SB}_r(A,H)$ and the result follows. □

As a result we need only speak of the strict boundary of H in G with respect to A. Example 8.3 tells us that, even for groups, the finiteness of the strict boundary may depend on the choice of the generating set A.

In an ideal world, whenever G is a finitely generated group and H a subgroup of G with finite index the strict right boundary of H in G would be finite. Unfortunately, this is not the case.

Proposition 8.19. *Let $N = \mathbb{Z}$, $H = \mathbb{Z}_2 \oplus \mathbb{Z}_2$ and let $G = N \oplus H = \mathbb{Z} \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$. Then for any finite generating set A of G the strict boundary of N in G, with respect to A, is infinite.*

Proof. Let A be a finite generating set for G. Since A generates G there is some $(x,y,z) \in A$ with $x \neq 0$ and some $(a,b,c) \in A$ such that $(b,c) \neq (y,z)$ and $(b,c) \neq (0,0)$. Define the following word in these generators:

$$w_{2k} = (a,b,c) + \underbrace{(x,y,z) + \ldots + (x,y,z)}_{2k} + (a,b,c).$$

We have $w_{2k} \in N$ since

$$w_{2k} = (2(a+kx), 2(b+ky), 2(c+kz)) = (2(a+kx), 0, 0) \in N.$$

We claim that $w_{2k} \in \mathcal{SWB}_r(A,N)$ and as a consequence that $w_{2k}\theta \in \mathcal{SB}_r(A,N) = \mathcal{SB}(A,N)$. First note that $(a,b,c) \notin N$ since $(b,c) \neq (0,0)$. Now let $l \in \{1, \ldots, 2k\}$ and consider the word:

$$v_l = (a,b,c) + \underbrace{(x,y,z) + \ldots + (x,y,z)}_{l} = (a+lx, b+ly, c+lz).$$

If l is even then $b+ly \equiv b \pmod 2$, $c+lz \equiv c \pmod 2$ and since $(b,c) \neq (0,0)$ it follows that $v_l \notin N$. On the other hand, if l is odd then $b+ly \equiv b+y \pmod 2$, $c+ly \equiv c+z \pmod 2$ and since $(b,c) \neq (y,z)$ then either $b \neq y$, which means that $b+y \equiv 1 \pmod 2$, or $c \neq z$, in which case $c+z \equiv 1 \pmod 2$. In either case we conclude that $v_l \notin N$.

It follows that for all $k \in \mathbb{N}$ the element $w_{2k}\theta \in \mathcal{SB}(A,H)$. □

In the example above N has index 4 in G but the strict boundary of N in G is infinite. One simple situation where we may obtain a positive result is given by the following lemma. Before stating the lemma we need the following definition.

Definition 8.20. Let G be a group and let H be a subgroup of G. Then

$$\mathrm{Core}(H) = \bigcap_{g \in G} gHg^{-1}.$$

$\mathrm{Core}(H)$ is a normal subgroup of G and it is well known that if H has finite index in G then so does $\mathrm{Core}(H)$.

Lemma 8.21. *Let G be a finitely generated group with H a subgroup of G. If $G/\mathrm{Core}(H) \cong \mathbb{Z}_m$ for some $m \in \mathbb{N}$ then H has a finite strict boundary in G.*

Proof. If $\mathrm{Core}(H) = G$ the result is trivial. Otherwise let C be a coset of $\mathrm{Core}(H)$ in G that corresponds to a generator of $G/\mathrm{Core}(H)$. It follows that $\langle C \rangle = G$ and, since G is finitely generated, there is a finite subset B of C that generates G. Then the strict right boundary of $\mathrm{Core}(H)$ in G with respect to B is finite since any product of generators of C of length m must belong to $\mathrm{Core}(H)$. □

The example given in Proposition 8.19 also answers the question of whether or not having a finite strict boundary is a transitive property for subgroups of groups. If we define $K = \mathbb{Z} \oplus \mathbb{Z}_2 \oplus \{0\}$ and let N and G be as in the proposition then it is clear that $N \leq K \leq G$. Moreover, the index of K in G is 2 and the index of N in K is 2. It follows from Lemma 8.21 that N has a finite strict boundary in K and that K has a finite strict boundary in G. Thus the property of having a finite strict boundary is not a transitive one, not even for groups. This leads to the following question.

> If G is a finitely presented group with H a subgroup of G with finite index, then can we always construct a finite chain of subgroups
>
> $$G = K_0 \geq K_1 \geq \ldots \geq K_m = H$$

such that K_i has finite strict boundary in K_{i-1} for all i?

We provide a partial answer to this question here. A group is *soluble* if it has a normal series such that each normal factor is abelian, and a group is *polycyclic* if it has a normal series such that each normal factor is cyclic. It is well known that a finite group if soluble if and only if it is polycyclic (i.e. every finite soluble group has a subnormal series with cyclic factors; see [80, 5.4.12.]).

Proposition 8.22. *Let G be a finitely generated group with N a normal subgroup of G. If G/N is a finite soluble group then there exists a finite chain*

$$G = K_l \geq \ldots \geq K_0 = N$$

such that the strict right boundary of K_i in K_{i+1} is finite for all i.

Proof. We prove the result by induction on the index of N in G. When $[G : N] = 2$ we have $G/N \cong \mathbb{Z}_2$ and the result follows from Lemma 8.21. Now suppose that the result holds for all groups G' with normal subgroups N' such that $[G' : N'] < [G : N]$ and G'/N' is finite and soluble. Since G/N is finite and soluble we can write:

$$G/N = L_0 \trianglerighteq L_1 \trianglerighteq \ldots \trianglerighteq L_k = \{1\}$$

where L_i/L_{i+1} is finite and cyclic for all i. By the correspondence theorem there exist groups:

$$G = K_0 \geq K_1 \geq \ldots \geq K_k = N$$

such that $K_i/N = L_i$ for all i. Also, since $L_1 \trianglelefteq L_0 = G/N$ it follows that $K_1 \trianglelefteq K_0 = G$. By construction we have $K_1/N = L_1$. Note that $[K_1 : N] < [G : N]$ and K_1/N is soluble so we can apply induction. By induction there is a chain:

$$K_1 = H_0 \geq H_1 \geq \ldots \geq H_l = N$$

such that H_{i+1} has a finite strict boundary in H_i for all i. By the third isomorphism theorem we have:

$$G/K_1 = \frac{G/N}{K_1/N} = L_0/L_1$$

where L_0/L_1 is a finite cyclic group. If follows from Lemma 8.21 that K_1 has a finite strict boundary in G. Therefore the chain:

$$G = G_0 \geq K_1 \geq H_1 \geq H_2 \geq \ldots \geq H_l = N$$

has the property that each group has a finite strict finite boundary in the previous one. This completes the inductive step and the proof of the proposition. □

We leave the general question as an open problem.

Open Problem 10. Let G be a finitely generated group with N a normal subgroup of G. Prove that if N has finite index in G there is a finite chain:

$$G = K_0 \geq K_1 \geq \ldots \geq K_r = N$$

such that K_{i+1} has a finite strict boundary in K_i for all i.

Subgroups of completely 0-simple semigroups

In Chapter 2 we undertook a fairly detailed analysis of the generating sets of finite completely 0-simple semigroups. It is obvious, using ideas from those chapters, that for $S = \mathcal{M}^0[G; I, \Lambda; P]$ to stand a chance of being finitely generated, both I and Λ must be finite. Furthermore, if G is finitely generated and I and Λ are finite then S is obviously finitely generated. What about the converses? These questions were considered in [6] where it was shown that S is finitely generated (resp. presented) if and only if I and Λ are finite and G is finitely generated (resp. presented). In this subsection we will recover one direction of this result as an application of Theorem 8.10.

Our approach is as follows. Given a finitely generated completely 0-simple semigroup $S = \mathcal{M}^0[G; I, \Lambda; P]$, where I and Λ are finite, we will show the existence of a subsemigroup T of S that is right unitary, has a finite strict left boundary in S, and has a subgroup H isomorphic to G that is left unitary in T and has a finite strict right boundary in T. It then follows from Theorem 8.10 that if S is finitely presented then T is finitely presented and so G is finitely presented.

Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ where $I = \{1, \ldots, m\}$, $\Lambda = \{1, \ldots, n\}$ and where P has been normalised (using Theorem 2.54) so that every non-zero entry in the first row and first column is equal to 1, and so that $p_{11} = 1$.

Proposition 8.23. *Let S be a completely 0-simple semigroup. Let R and L be a non-zero $\mathcal{R}$- and $\mathcal{L}$-class, respectively, of S.*

(i) The union of group $\mathcal{H}$-classes of S that intersect R is a right unitary subsemigroup of S.

(ii) The union of group $\mathcal{H}$-classes of S that intersect L is a left unitary subsemigroup of S.

Proof. We prove part (i) only. Part (ii) may be proven using a dual argument. Let $s \in S \setminus T$ and let $t \in T$. There are three possibilities to consider.

Case 1: $s = 0$. In this case $st = 0$ which is not in T.

Case 2: $s \neq 0$ *and* $L_s \cap T = \varnothing$. In this case $st = 0$ which is not in T.

Case 3: $s \neq 0$ *and* $L_s \cap T \neq \varnothing$. Then st belongs to the same $\mathcal{R}$-class as s which means that $st \notin T$.

It follows that T is right unitary in S. ☐

By exactly the same argument we have:

Proposition 8.24. *Let S be a completely simple semigroup. Let R and L be a non-zero $\mathcal{R}$- and $\mathcal{L}$-class, respectively, of S. Then R and L are right and left unitary subsemigroups, respectively, of S.* ☐

We now want to show that these subsemigroups have finite strict left or right boundaries in S. In order to do this we will give a method for transforming an arbitrary finite generating set of S into one with respect to which the subsemigroup R (or L) has a finite strict boundary.

Definition 8.25. Define $f : S \to \mathcal{P}(S)$ by

$$f(\ (i, g, \lambda)\ ) = \begin{cases} \{(i, g, \lambda)\} & \text{if } i = 1 \\ \{(1, g, \lambda), (i, 1, 1)\} & \text{otherwise.} \end{cases}$$

Lemma 8.26. *If $A \subseteq S$ generates S and $p_{11} = 1$ then $f(A) = \bigcup_{a \in A} f(a)$ generates S.*

Proof. Since $(i, 1, 1)(1, g, \lambda) = (i, 1g, \lambda) = (i, g, \lambda)$ it follows that $A \subseteq \langle f(A) \rangle$ and so $\langle f(A) \rangle \supseteq \langle A \rangle = S$. ☐

Lemma 8.27. *Let A be a finite generating set for S. Let $T = \{(1, g, \lambda) : g \in G, p_{\lambda 1} = 1\}$. Then the strict left boundary of T in S with respect to $f(A)$ is finite.*

Proof. It is immediate from the definitions that every element in $\mathcal{SB}_l(f(A),T)$ has the form

$$(1,g_k,\lambda)(i_{k-1},1,1)(i_{k-2},1,1)\dots(i_1,1,1)$$

where $i_j \neq 1$ for all j. Now we have

$$(1,g_k,\lambda)(i_{k-1},1,1)\dots(i_1,1,1) = (1,g_k p_{\lambda i_{k-1}} 1 p_{1 i_{k-2}} \dots p_{1 i_1} 1, 1) = (1, g_k p_{\lambda i_{k-1}}, 1).$$

We conclude that $\mathcal{SB}_l(f(A),T) \subseteq \{(1, g p_{\lambda,i}, 1) : g \in m(f(A)), i \in I, \lambda \in \Lambda\}$ which is a finite since $f(A)$ is finite and I and Λ are both finite. Here $m(f(A))$ denotes the set of middle components of the set $f(A)$ which is a subset of G. □

Now we do the same thing but for H_{11} inside T. first we define the dual of the function f.

Definition 8.28. Define $f' : S \to \mathcal{P}(S)$ by

$$f'(\ (i,g,\lambda)\ ) = \begin{cases} \{(i,g,\lambda)\} & \text{if } \lambda = 1 \\ \{(1,g,\lambda),(1,1,\lambda)\} & \text{otherwise.} \end{cases}$$

In the same way as in Lemma 8.26 if A generates T then $f'(A)$ generates T. Let H_{11} be the $\mathcal{H}$-class $H_{11} = \{(1,g,1) : g \in G\}$. It is a group since $p_{11} = 1$ by assumption.

Lemma 8.29. *The subsemigroup H_{11} is left unitary in T and if A is any finite generating set for T then H_{11} has a finite strict right boundary in T with respect to the generating set $f'(A)$.*

Proof. For the first part let $h \in H_{11}$ and let $t \in T \setminus H_{11}$. Then ht is in the $\mathcal{L}$-class of t in T and so does not belong to H_{11}.

For the second part consider an arbitrary element of $\mathcal{SB}_r(f'(A), H_{11})$ in T. It can be written as

$$(1,1,\lambda_1)\dots(1,1,\lambda_{k-1})(1,g_k,1)$$

where $\lambda_j \neq 1$ for all j. Now we have

$$(1,1,\lambda_1)\dots(1,1,\lambda_{k-1})(1,g_k,1) = (1, p_{\lambda_1 1} p_{\lambda_2 1} \dots p_{\lambda_{k-1} 1} g_k, 1) = (1,g_k,1).$$

Since $f'(A)$ is finite it follows that the strict right boundary of H_{11} in T is finite. □

From these results we can conclude the following.

Theorem 8.30. *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 is finitely generated (resp. finitely presented) then G is finitely generated (resp. finitely presented).*

Proof. 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 by Proposition 8.23 and Lemma 8.27 the subsemigroup T is right unitary in S and has a finite strict left boundary in S. By Lemma 8.29 the subsemigroup G is left unitary in T and has a finite strict right boundary in T. The result now follows by applying Theorem 8.10. □

We conclude this chapter by mentioning that in [83] a much more general result than the one above was proven. It was shown that if S is a regular semigroup with finitely many $\mathcal{R}$- and $\mathcal{L}$-classes then S is finitely presented if and only if all of its maximal subgroups are finitely presented.

Bibliography

[1] *Proceedings of the International Conference on Semigroups and Languages In Honour of the 65 th birthday of Donald B. McAlister*. World Scientific, (to appear).

[2] A. Ja. Aĭzenštat. Defining relations in symmetric semigroups. *Leningrad. Gos. Ped. Inst. Uč. Zap.*, 166:121–142, 1958.

[3] A. Ja. Aĭzenštat. Defining relations of finite symmetric semigroups. *Mat. Sb. N.S.*, 45 (87):261–280, 1958.

[4] A. Ja. Aĭzenštat. The defining relations of the endomorphism semigroup of a finite linearly ordered set. *Sibirsk. Mat. Ž.*, 3:161–169, 1962.

[5] Isabel M. Araújo, Mário J. J. Branco, Vitor H. Fernandes, Gracinda M. S. Gomes, and N. Ruškuc. On generators and relations for unions of semigroups. *Semigroup Forum*, 63(1):49–62, 2001.

[6] H. Ayik and N. Ruškuc. Generators and relations of Rees matrix semigroups. *Proc. Edinburgh Math. Soc. (2)*, 42(3):481–495, 1999.

[7] Andrew Baker. *Matrix groups*. Springer Undergraduate Mathematics Series. Springer-Verlag London Ltd., London, 2002. An introduction to Lie group theory.

[8] Jørgen Bang-Jensen and Gregory Gutin. *Digraphs*. Springer Monographs in Mathematics. Springer-Verlag London Ltd., London, 2001. Theory, algorithms and applications.

[9] Gilbert Baumslag, Martin R. Bridson, Derek F. Holt, and Charles F. Miller, III. Finite presentation of fibre products of metabelian groups. *J. Pure Appl. Algebra*, 181(1):15–22, 2003.

[10] Béla Bollobás. *Graph theory*, volume 63 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1979. An introductory course.

[11] A. E. Brouwer, A. M. Cohen, and A. Neumaier. *Distance-regular graphs*, volume 18 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1989.

[12] S. Bulman-Fleming. Regularity and products of idempotents in endomorphism monoids of projective acts. *Mathematika*, 42(2):354–367, 1995.

[13] J. M. Burns and B. Goldsmith. Maximal order abelian subgroups of symmetric groups. *Bull. London Math. Soc.*, 21(1):70–72, 1989.

[14] P.. Cameron and C. Szabó. Independence algebras. *J. London Math. Soc. (2)*, 61(2):321–334, 2000.

[15] Peter J. Cameron. *Combinatorics: topics, techniques, algorithms*. Cambridge University Press, Cambridge, 1994.

[16] Peter J. Cameron, Ron Solomon, and Alexandre Turull. Chains of subgroups in symmetric groups. *J. Algebra*, 127(2):340–352, 1989.

[17] C. M. Campbell, E. F. Robertson, N. Ruškuc, and R. M. Thomas. Reidemeister-Schreier type rewriting for semigroups. *Semigroup Forum*, 51(1):47–62, 1995.

[18] C. M. Campbell, E. F. Robertson, N. Ruškuc, and R. M. Thomas. On subsemigroups of finitely presented semigroups. *J. Algebra*, 180(1):1–21, 1996.

[19] C. M. Campbell, E. F. Robertson, N. Ruškuc, and R. M. Thomas. Presentations for subsemigroups—applications to ideals of semigroups. *J. Pure Appl. Algebra*, 124(1-3):47–64, 1998.

[20] W. Edwin Clark. Matching subspaces with complements in finite vector spaces. *Bull. Inst. Combin. Appl.*, 6:33–38, 1992.

[21] 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.

[22] Peter M. Curran. Subgroups of finite index in certain classes of finitely presented groups. *J. Algebra*, 122(1):118–129, 1989.

[23] Francesca Dalla Volta and Andrea Lucchini. Finite groups that need more generators than any proper quotient. *J. Austral. Math. Soc. Ser. A*, 64(1):82–91, 1998.

[24] R. Dawlings. On idempotent affine mappings. *Proc. Roy. Soc. Edinburgh Sect. A*, 93(3-4):345–348, 1982/83.

[25] R. J. H. Dawlings. Products of idempotents in the semigroup of singular endomorphisms of a finite-dimensional vector space. *Proc. Roy. Soc. Edinburgh Sect. A*, 91(1-2):123–133, 1981/82.

[26] R. J. H. Dawlings. Sets of idempotents that generate the semigroup of singular endomorphisms of a finite-dimensional vector space. *Proc. Edinburgh Math. Soc. (2)*, 25(2):133–139, 1982.

[27] Klaus Denecke and Shelly L. Wismath. *Universal algebra and applications in theoretical computer science*. Chapman & Hall/CRC, Boca Raton, FL, 2002.

[28] M. N. Ellingham and J. D. Horton. Non-Hamiltonian 3-connected cubic bipartite graphs. *J. Combin. Theory Ser. B*, 34(3):350–353, 1983.

[29] J. A. Erdos. On products of idempotent matrices. *Glasgow Math. J.*, 8:118–122, 1967.

[30] J. Fountain and A. Lewin. Products of idempotent endomorphisms of an independence algebra of finite rank. *Proc. Edinburgh Math. Soc. (2)*, 35(3):493–500, 1992.

[31] John Fountain. Products of idempotent integer matrices. *Math. Proc. Cambridge Philos. Soc.*, 110(3):431–441, 1991.

[32] John Fountain and Andrew Lewin. Products of idempotent endomorphisms of an independence algebra of infinite rank. *Math. Proc. Cambridge Philos. Soc.*, 114(2):303–319, 1993.

[33] John B. Fountain. The depth of the semigroup of balanced endomorphisms. *Mathematika*, 41(1):199–208, 1994.

[34] G. U. Garba. Idempotents in partial transformation semigroups. *Proc. Roy. Soc. Edinburgh Sect. A*, 116(3-4):359–366, 1990.

[35] G. U. Garba. On the nilpotent rank of partial transformation semigroups. *Portugal. Math.*, 51(2):163–172, 1994.

[36] G. U. Garba. On the nilpotent ranks of certain semigroups of transformations. *Glasgow Math. J.*, 36(1):1–9, 1994.

[37] G. M. S. Gomes and J. M. Howie. On the ranks of certain finite semigroups of transformations. *Math. Proc. Cambridge Philos. Soc.*, 101(3):395–403, 1987.

[38] Gracinda M. S. Gomes and John M. Howie. On the ranks of certain semigroups of order-preserving transformations. *Semigroup Forum*, 45(3):272–282, 1992.

[39] V. Gould. Independence algebras. *Algebra Universalis*, 33(3):294–318, 1995.

[40] N. Graham, R. Graham, and J. Rhodes. Maximal subsemigroups of finite semigroups. *J. Combinatorial Theory*, 4:203–209, 1968.

[41] R. L. Graham. On finite 0-simple semigroups and graph theory. *Math. Systems Theory*, 2:325–339, 1968.

[42] R. Gray. Hall's condition and idempotent rank of ideals of endomorphism monoids. *submitted.*

[43] R. Gray. Idempotent rank in endomorphism monoids of finite independence algebras. *submitted.*

[44] R. Gray and J. D. Mitchell. Largest subsemigroups of the full transformation monoid. *submitted.*

[45] R. Gray and N. Ruškuc. Generators and relations for subsemigroups via boundaries in cayley graphs. *submitted.*

[46] R. Gray and N. Ruškuc. Generating sets of completely 0-simple semigroups. *Comm. Algebra*, 33:4657–4678, 2005.

[47] M. A. Grechkoseeva. On minimal permutation representations of classical simple groups. *Sibirsk. Mat. Zh.*, 44(3):560–586, 2003.

[48] J. A. Green. On the structure of semigroups. *Ann. of Math. (2)*, 54:163–172, 1951.

[49] P.-A. Grillet. *Semigroups*, volume 193 of *Monographs and Textbooks in Pure and Applied Mathematics.* Marcel Dekker Inc., New York, 1995. An introduction to the structure theory.

[50] Philip Hall. *The collected works of Philip Hall.* Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1988.

[51] Frank Harary. *Graph theory.* Addison-Wesley Publishing Co., Reading, Mass.-Menlo Park, Calif.-London, 1969.

[52] Peter M. Higgins. *Techniques of semigroup theory.* Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1992. With a foreword by G. B. Preston.

[53] Michael Hoffmann, Richard M. Thomas, and Nik Ruškuc. Automatic semigroups with subsemigroups of finite Rees index. *Internat. J. Algebra Comput.*, 12(3):463–476, 2002.

[54] C. H. Houghton. Completely 0-simple semigroups and their associated graphs and groups. *Semigroup Forum*, 14(1):41–67, 1977.

[55] J. M. Howie. The subsemigroup generated by the idempotents of a full transformation semigroup. *J. London Math. Soc.*, 41:707–716, 1966.

[56] J. M. Howie. Idempotents in completely 0-simple semigroups. *Glasgow Math. J.*, 19(2):109–113, 1978.

[57] J. M. Howie. *Fundamentals of semigroup theory.* Academic Press [Harcourt Brace Jovanovich Publishers], London, 1995. L.M.S. Monographs, No. 7.

[58] John M. Howie. Why study semigroups? *Math. Chronicle*, 16:1–14, 1987.

[59] John M. Howie and Robert B. McFadden. Idempotent rank in finite full transformation semigroups. *Proc. Roy. Soc. Edinburgh Sect. A*, 114(3-4):161–167, 1990.

[60] John M. Howie, N. Ruškuc, and P. M. Higgins. On relative ranks of full transformation semigroups. *Comm. Algebra*, 26(3):733–748, 1998.

[61] Andrzej Jura. Determining ideals of a given finite index in a finitely presented semigroup. *Demonstratio Math.*, 11(3):813–827, 1978.

[62] Andrzej Jura. Some remarks on nonexistence of an algorithm for finding all ideals of a given finite index in a finitely presented semigroup. *Demonstratio Math.*, 13(2):573–578, 1980.

[63] L. G. Kovács and Cheryl E. Praeger. On minimal faithful permutation representations of finite groups. *Bull. Austral. Math. Soc.*, 62(2):311–317, 2000.

[64] Gérard Lallement. *Semigroups and combinatorial applications.* John Wiley & Sons, New York-Chichester-Brisbane, 1979. Pure and Applied Mathematics, A Wiley-Interscience Publication.

[65] I. Levi and R. B. McFadden. S_n-normal semigroups. *Proc. Edinburgh Math. Soc. (2)*, 37(3):471–476, 1994.

[66] I. Levi and S. Seif. Combinatorial techniques for determining rank and idempotent rank of certain finite semigroups. *Proc. Edinb. Math. Soc. (2)*, 45(3):617–630, 2002.

[67] I. Levi and S. Seif. Constructive techniques for labeling constant weight Gray codes with applications to minimal generating sets of semigroups. *J. Algebra*, 266(1):189–219, 2003.

[68] I. Levi and S. Seif. Counting techniques to label constant weight Gray codes with links to minimal generating sets of semigroups. *J. Algebra*, 266(1):220–238, 2003.

[69] Inessa Levi and G. R. Wood. On maximal subsemigroups of Baer-Levi semigroups. *Semigroup Forum*, 30(1):99–102, 1984.

[70] Martin W. Liebeck, Cheryl E. Praeger, and Jan Saxl. A classification of the maximal subgroups of the finite alternating and symmetric groups. *J. Algebra*, 111(2):365–383, 1987.

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

[72] Wilhelm Magnus, Abraham Karrass, and Donald Solitar. *Combinatorial group theory.* Dover Publications Inc., New York, revised edition, 1976. Presentations of groups in terms of generators and relations.

[73] Al. A. Markov. On finitely generated subsemigroups of a free semigroup. *Semigroup Forum*, 3(3):251–258, 1971/1972.

[74] R. N. McKenzie, G. F. McNulty, and W. F. Taylor. *Algebras, lattices, varieties. Vol. I.* The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, 1987.

[75] J. Moon and L. Moser. On Hamiltonian bipartite graphs. *Israel J. Math.*, 1:163–165, 1963.

[76] W. Narkiewicz. Independence in a certain class of abstract algebras. *Fund. Math.*, 50:333–340, 1961/1962.

[77] Katsumi Nomizu. *Fundamentals of linear algebra*. McGraw-Hill Book Co., New York, 1966.

[78] Norman R. Reilly. Maximal inverse subsemigroups of $T_X^{(1)}$. *Semigroup Forum*, 15(4):319–326, 1977/78.

[79] M. A. Reynolds and R. P. Sullivan. Products of idempotent linear transformations. *Proc. Roy. Soc. Edinburgh Sect. A*, 100(1-2):123–138, 1985.

[80] Derek J. S. Robinson. *A course in the theory of groups*, volume 80 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1996.

[81] N. Ruškuc. On the rank of completely 0-simple semigroups. *Math. Proc. Cambridge Philos. Soc.*, 116(2):325–338, 1994.

[82] N. Ruškuc. On large subsemigroups and finiteness conditions of semigroups. *Proc. London Math. Soc. (3)*, 76(2):383–405, 1998.

[83] N. Ruškuc. Presentations for subgroups of monoids. *J. Algebra*, 220(1):365–380, 1999.

[84] N. Ruškuc. On finite presentability of monoids and their Schützenberger groups. *Pacific J. Math.*, 195(2):487–509, 2000.

[85] Carla D. Savage and Peter Winkler. Monotone Gray codes and the middle levels problem. *J. Combin. Theory Ser. A*, 70(2):230–248, 1995.

[86] Marcel Paul Schützenberger. Une théorie algébrique du codage. *C. R. Acad. Sci. Paris*, 242:862–864, 1956.

[87] Štefan Schwarz. On some semigroups in combinatorics. In *Mini-Conference on Semigroup Theory (József Attila Univ., Szeged, 1972)*, pages 24–31. József Attila Univ., Szeged, 1973.

[88] Craig C. Squier. Word problems and a homological finiteness condition for monoids. *J. Pure Appl. Algebra*, 49(1-2):201–217, 1987.

[89] Craig C. Squier, Friedrich Otto, and Yuji Kobayashi. A finiteness condition for rewriting systems. *Theoret. Comput. Sci.*, 131(2):271–294, 1994.

[90] R. Strebel. Subgroups of finitely presented centre-by-metabelian groups. *J. London Math. Soc. (2)*, 28(3):481–491, 1983.

[91] R. P. Sullivan. Semigroups generated by nilpotent transformations. *J. Algebra*, 110(2):324–343, 1987.

[92] K. Urbanik. Linear independence in abstract algebras. *Colloq. Math.*, 14:233–255, 1966.

[93] Jing Wang. Finite complete rewriting systems and finite derivation type for small extensions of monoids. *J. Algebra*, 204(2):493–503, 1998.

[94] William C. Waterhouse. Two generators for the general linear groups over finite fields. *Linear and Multilinear Algebra*, 24(4):227–230, 1989.

[95] Xiuliang Yang. Non-group ranks in finite full transformation semigroups. *Semigroup Forum*, 57(1):42–47, 1998.

[96] Xiuliang Yang. On the nilpotent ranks of the principal factors of order-preserving transformation semigroups. *Semigroup Forum*, 57(3):331–340, 1998.

[97] Xiuliang Yang. On the nilpotent ranks of the principle factors of certain semigroups of partial transformations. *Comm. Algebra*, 26(5):1445–1455, 1998.

[98] Xiuliang Yang. A classification of maximal inverse subsemigroups of the finite symmetric inverse semigroups. *Comm. Algebra*, 27(8):4089–4096, 1999.

[99] Xiuliang Yang. A classification of maximal subsemigroups of finite order-preserving transformation semigroups. *Comm. Algebra*, 28(3):1503–1513, 2000.

[100] Xiuliang Yang. Maximal subsemigroups of the finite singular transformation semigroup. *Comm. Algebra*, 29(3):1175–1182, 2001.

[101] Louxin Zhang. Conjugacy in special monoids. *J. Algebra*, 143(2):487–497, 1991.

[102] Louxin Zhang. Applying rewriting methods to special monoids. *Math. Proc. Cambridge Philos. Soc.*, 112(3):495–505, 1992.