Cosets in Inverse Semigroups and Inverse Subsemigroups of Finite Index

Amal AlAli

Submitted for the degree of

Doctor of Philosophy

Heriot-Watt University

Department of Mathematics,

School of Mathematical and Computer Sciences.

July 2016

Abstract

The index of a subgroup of a group counts the number of cosets of that subgroup. A subgroup of finite index often shares structural properties with the group, and the existence of a subgroup of finite index with some particular property can therefore imply useful structural information for the overgroup. Although a developed theory of cosets in inverse semigroups exists, it is defined only for closed inverse subsemigroups, and the structural correspondences between an inverse semigroup and a closed inverse subsemigroup of finte index are much weaker than in the group case. Nevertheless, many aspects of this theory remain of interest, and some of them are addressed in this thesis.

We study the basic theory of cosets in inverse semigroups, including an index formula for chains of subgroups and an analogue of M. Hall's Theorem on counting subgroups of finite index in finitely generated groups. We then look at specific examples, classifying the finite index inverse subsemigroups in polycyclic monoids and in graph inverse semigroups. Finally, we look at the connection between the properties of finite generation and having finte index: these were shown to be equivalent for free inverse monoids by Margolis and Meakin.

Acknowledgements

First and foremost, I would like to express my deep thanks and gratitude to my supervisor Prof. Nick Gilbert, whose expertise, guidance, understanding, and patience, added considerably to this thesis. This work could not have been completed without his unwavering support and encouragement.

I would also like to thank my parents, to whom I owe my life for their constant love, encouragement, and blessings.

I also acknowledge with deep thanks my husband, Waleed. Without his love, patience, support and belief in me, I would not have finished this thesis.

Thank you to my sister and brothers for all their support and encouragement they provided me through my entire life.

Very special thanks to my children for their complete and unconditional love that carries me through always.

Finally, I am grateful for Princess Norah University, and the Ministry of Education; Kingdom of Saudi Arabia for their financial support.

ACADEMIC REGISTRY
Research Thesis Submission

Name:	AMAL ALALI		
School/PGI:	MACS		
Version: *(i.e. First, Resubmission, Final)*	FIRST	Degree Sought (Award **and** Subject area)	PhD Mathematics

Declaration

In accordance with the appropriate regulations I hereby submit my thesis and I declare that:

1) the thesis embodies the results of my own work and has been composed by myself
2) where appropriate, I have made acknowledgement of the work of others and have made reference to work carried out in collaboration with other persons
3) the thesis is the correct version of the thesis for submission and is the same version as any electronic versions submitted*.
4) my thesis for the award referred to, deposited in the Heriot-Watt University Library, should be made available for loan or photocopying and be available via the Institutional Repository, subject to such conditions as the Librarian may require
5) I understand that as a student of the University I am required to abide by the Regulations of the University and to conform to its discipline.

* *Please note that it is the responsibility of the candidate to ensure that the correct version of the thesis is submitted.*

Signature of Candidate:		Date:	

Submission

Submitted By *(name in capitals):*	
Signature of Individual Submitting:	
Date Submitted:	

For Completion in the Student Service Centre (SSC)

Received in the SSC by *(name in capitals):*			
Method of Submission *(Handed in to SSC; posted through internal/external mail):*			
*E-thesis Submitted (**mandatory for final theses)***			
Signature:		Date:	

Please note this form should bound into the submitted thesis.

Updated February 2008, November 2008, February 2009, January 2011

Contents

Chapter 1

Introduction

It has long been recognized that the origin and development of the theory of inverse semigroups are due to the substantial contributions of two prime foci: the Soviet and Western schools, led by Wagner and Schein in the Soviet school, and Preston and Munn in the school of the West. The concept of inverse semigroups was defined first by Wagner [31] in 1952 and called "generalized groups" and he and some of his followers have used this term ever since. The term "inverse semi-groups" was introduced by Preston [21] who independently discovered this class of semigroups in 1954.

Inverse semigroups are the mathematical structures which can provide a full description of the concept of partial symmetries. In this sense, inverse semigroups are the generalizations of the idea of transforming symmetries in natural forms into the mathematical precise notion of groups. There has been a large body of literature concerning the structure of inverse semigroups. The reason for this is probably due to the simple and beautifully constructed theorems for different classes of inverse semigroups.

Therefore, it is not surprising that the subject has attracted the attention of many researchers. Many structure theorems and concepts for various classes of semigroups have been formulated through the years which give the subject a broad spectrum of various applications and connections with other branches of mathematics. In fact, inverse semigroups recently have wide applications in many aspects of modern sciences including: geometric group theory, combinatorial group theory, model theory, linear logic, solid state physics and C^*- algebras.

It is worth noting that the similarities between groups and inverse semigroups are not as substantial as it might appear at first sight. However, there is a significant analogy between the two theories. Inverse semigroups represent the concept of sets of partial symmetries just as groups play the same role for groups of symmetries.

Consequently, there has been a tendency among many researchers in the field of semigroups to model structural theorems on those in group theory. Even though this approach has had a limited success, groups still play a significant role in many important theorems for different classes of inverse semigroups. In the present research, the process of generalizing some notions and group theoretical approaches to semigroups, in particular inverse semigroups, is considered in a way that is similar in spirit to the group-theoretic formalization.

Nevertheless, there are some significant differences between the two theories. Among them is the natural partial order that exists in each inverse semigroup. Moreover, the basic theoretical structure of semigroups is very different from that of groups: for example, a congruence on a semigroup is not generally determined by one congruence class, as is the case in groups.

For further background information on the general theory of inverse semigroups, see [11, 14, 20]

1.1 Regular and Inverse Semigroups

A *semigroup* is defined as a non-empty set S on which an associative *binary operation* is defined, by which we mean a map $f : S \times S \longrightarrow S$ written $(x, y) \mapsto xy$ such that for all $x, y, z \in S$, we have $(xy)z = x(yz)$. If a semigroup has an identity element $1 \in S$, then S is called a *monoid*. A semigroup has at most one identity element.

If a semigroup S has no identity element, then we can adjoin an extra element 1 to S to form a monoid. We define

$$1s = s1 = s \quad \text{for all } s \in S, \text{ and } 11 = 1,$$

and then $S \cup \{1\}$ becomes a monoid. Thus we now can define

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

If a semigroup S has the property that, for all $x, y \in S,\ xy = yx,$ we shall say that S is a *commutative semigroup.*

An element x of a semigroup S is called *regular,* if there exists $y \in S$ such that

$$x = x\,y\,x.$$

The semigroup S is called *a regular semigroup* if all its elements are regular.

A semigroup S is called an *inverse semigroup*, if for each $s \in S$ there is a unique element $s^{-1} \in S$ such that

$$s\,s^{-1}\,s = s \quad \textit{and} \quad s^{-1}\,s\,s^{-1} = s^{-1}\,.$$

An element e of an inverse semigroup S is an *idempotent* if $e^2 = e$. The set of idempotents of an inverse semigroup S is a commutative inverse subsemigroup of S. We will consistently denote the *set of idempotents* of an inverse semigroup S by $E(S)$.

A non-empty subset T of S is called a *subsemigroup* of S, if it is closed with respect to multiplication, that is

$$xy \in T \quad \text{for all } x, y \in T\,.$$

Furthermore, if T is closed under taking inverses, then it is called an *inverse subsemigroup* of S. For any idempotent $e \in E(S)$ the subset $eSe = \{ese : s \in S\}$ is an inverse subsemigroup of S, called a *local* subsemigroup. An inverse subsemigroup T of S is said to be *full* if $E(T) = E(S)$.

A non-empty subset A of S is called a *left ideal* if $SA \subseteq A$, a *right ideal* if $AS \subseteq A$, and a (two-sided) *ideal* if it is both a left and a right ideal. Certainly, every ideal (whether right, left or two-sided) is a subsemigroup, but the converse is not true.
A map $\phi : S \longrightarrow T$, where $(S, .)$ and $(T, .)$ are semigroups, is called a *morphism* if

$$(xy)\phi = (x\phi)\,(y\phi) \quad \text{for all } x, y \in S.$$

If S and T are monoids, with identity elements $1_S, 1_T$ respectively, then ϕ will be called a *monoid morphism* if it has the additional property $1_S\,\phi = 1_T$.

1.2 Properties of Inverses

We record some useful results drawn from of [14, Section 1.4].

Proposition 1.2.1. *Let S be an inverse semigroup.*

(1) For any $s \in S$, both $s^{-1}s$ and ss^{-1} are idempotents and $s(s^{-1}s) = s$ and $(ss^{-1})s = s$.

(2) $(s^{-1})^{-1} = s$ for every $s \in S$.

(3) For any idempotent e in S and any $s \in S$, the element $s^{-1}es$ is an idempotent.

(4) *If* e *is an idempotent in* S, *then* $e^{-1} = e$.

(5) $(s_1 \cdots s_n)^{-1} = {s_n}^{-1} \cdots {s_1}^{-1}$ *for all* $s_1, \ldots, s_n \in S$ *where* $n \geqslant 2$.

Lemma 1.2.2. *Let* S *be an inverse semigroup.*

(1) *For every idempotent* e *and element* s *there exists an idempotent* f *such that* $e\,s = s\,f$.

(2) *For every idempotent* e *and element* s *there exists an idempotent* f *such that* $s\,e = f\,s$.

Proof. We prove the first part and the second is similar. Write $f = s^{-1}es$, which is an idempotent by Proposition 1.2.2(3). Then

$$sf = s(s^{-1}\,es) = (ss^{-1})es = e(ss^{-1})s = es\,.$$

□

Proposition 1.2.3. *Groups are precisely the inverse semigroups with exactly one idempotent.*

1.3 The Natural Partial Order

We will start by defining what we mean by a partial order relation, and then will see that it is very natural to define such a relation on inverse semigroups.

Definition A relation $\leqslant$ is a *partial order* on a set X, if it satisfies the following conditions:

1. Reflexivity: $x \leqslant x$ for all $x \in X$.
2. Antisymmetry: $x \leqslant y$ and $y \leqslant x$ imply that $x = y$.
3. Transitivity: $x \leqslant y$ and $y \leqslant z$ imply that $x \leqslant z$.

The set X is *partially ordered set* or a *poset*. A subset Q of P is said to be an *order ideal* if $x \leqslant y \in Q$ implies that $x \in Q$.

Every inverse semigroup S comes equipped with a *natural partial order* defined by:

$$s \leqslant t \quad \textit{if and only if} \quad s = t\,e \quad \textit{for some} \ \ e \in E(S).$$

The following theorem is of great importance for our purpose. In fact, the use of the argument it contains at several different points in this thesis, justifies its explicit statement and detailed proof (see [14, Lemma 1.4.6]).

Theorem 1.3.1. *Let S be an inverse semigroup with semilattice E of idempotents, and let $s, t \in S$. The following statements are equivalent:*

(1) $s \leqslant t$;

(2) $s = ft$ *for some idempotent* $f \in E$;

(3) $s^{-1} \leqslant t^{-1}$;

(4) $s = ss^{-1}t$;

(5) $s = ts^{-1}s$;

(6) $ss^{-1} = ts^{-1}$;

(7) $s = s\,t^{-1}\,s$;

(8) $ss^{-1} = st^{-1}$;

(9) $s^{-1}s = t^{-1}s$;

(10) $s^{-1}s = s^{-1}t$;

Proof. $(1) \Rightarrow (2)$. Let $s = t\,e$. Then $s = f\,t$, for some idempotent f by Lemma 1.2.2.

$(2) \Rightarrow (3)$. Let $s = f\,t$, for some idempotent f. Then by taking inverses, $s^{-1} = t^{-1}\,f$. By definition $s^{-1} \leqslant t^{-1}$.

$(3) \Rightarrow (4)$. Let $s^{-1} \leqslant t^{-1}$. Then $s^{-1} = t^{-1}e$, for some idempotent e. Then by taking inverses $s = e\,t$. But $e\,s = s$, and so $e\,s\,s^{-1} = s\,s^{-1}$. Thus $s = s\,s^{-1}\,t$.

$(4) \Rightarrow (5)$. Let $s = s\,s^{-1}\,t$. Then $s = t\,e$, for some idempotent e by Lemma 1.2.2. But $s\,e = s$, and so $s^{-1}\,s\,e = s^{-1}\,s$. Thus $s = t\,s^{-1}\,s$.

$(5) \Rightarrow (1)$. Let $s = t\,s^{-1}\,s$. Then $s = t\,e$, for some idempotent e. Hence by definition $s = t\,e$.

$(2) \Rightarrow (6)$. Let $s = f\,t$, where f is an idempotent. Thus $s\,s^{-1} = f\,t\,t^{-1}\,f = t\,t^{-1}\,f = t\,(f\,t)^{-1} = t\,s^{-1}$.

$(6) \Rightarrow (7)$. Let $ss^{-1} = ts^{-1}$. Then $s^{-1} = s^{-1}\,s\,s^{-1} = s^{-1}\,t\,s^{-1}$. By taking inverses for both sides, we obtain $s = s\,t^{-1}\,s$.

$(7) \Rightarrow (1)$. Let $s = s\,t^{-1}\,s$. Then $(st^{-1})^2 = st^{-1}$ and $(t^{-1}s)^2 = t^{-1}s$ and so $s = e\,s$ and $s = s\,f$, where $e = s\,t^{-1}$ and $f = t^{-1}\,s$ are idempotents. Now $e = s\,t^{-1} = s\,t^{-1}\,t\,t^{-1} = e\,t\,t^{-1} = t\,t^{-1}\,e$, since idempotents commute. But $s = e\,s = t\,t^{-1}\,e\,s = t\,t^{-1}\,s = t\,f$. Hence $s \leqslant t$.

$(7) \Rightarrow (8)$. Let $s = s\,t^{-1}\,s$. Then $ss^{-1} = s\,t^{-1}\,s\,s^{-1}$. Thus $(st^{-1})^2 = st^{-1}$, and so $st^{-1} = e$, where e is an idempotent. Also $s\,s^{-1}\,s = s$, and $ss^{-1}\,st^{-1} = st^{-1}$. Then $s\,s^{-1}\,e = e$, and $ss^{-1} = f$, where f is an idempotent. Hence, we get $fe = e$, and therefore $st^{-1} = f\,e$. By (7), $ss^{-1} = st^{-1}\,ss^{-1} = f\,e\,f = e\,f = e$, since idempotents commute. Since $ss^{-1} = e$, and $st^{-1} = e$, we obtain $ss^{-1} = st^{-1}$.

$(8) \Rightarrow (1)$. Let $ss^{-1} = st^{-1}\,s$. Then $s^{-1}\,ss^{-1} = s^{-1}\,st^{-1}$, and so $s^{-1} = s^{-1}\,st^{-1}$. Taking inverses, we obtain $s = ts^{-1}\,s$. But $s^{-1}s$ is an idempotent, and therefore $s = te$.

$(5) \Rightarrow (9)$. Let $s = t\,s^{-1}\,s$. Then $s^{-1}s = s^{-1}\,t\,s^{-1}s$. Thus $s^{-1}t = e$, where e is an idempotent, and so $t^{-1}s = e^{-1} = e$. Also $t^{-1}s = e$. We know $s^{-1}\,ss^{-1} = s^{-1}$, and $s^{-1}\,ss^{-1}t = s^{-1}\,t$. Then $s^{-1}\,s\,e = e$, and $s^{-1}\,s = f$, where f is an idempotent. Hence, we get $fe = e$, and therefore $s^{-1}t = f\,e$. By (5), $s^{-1}\,s = s^{-1}\,ts^{-1}\,s = f\,e\,f = e\,f = e$, since idempotents commute. Now $s^{-1}\,s = e$, and $t^{-1}\,s = e$, and so we obtain $s^{-1}\,s = t^{-1}\,s$.

$(9) \Rightarrow (10)$. Let $s^{-1}s = t^{-1}\,s$. Then by taking inverses, we get $s^{-1}s = s^{-1}\,t$.

$(10) \Rightarrow (2)$. Let $s^{-1}s = s^{-1}t$. Then $ss^{-1}s = ss^{-1}t$. But ss^{-1} is an idempotent, and $ss^{-1}s = s$. Therefore $s = ft$. $\square$

The following result presents a number of different properties of the natural partial order.

Proposition 1.3.2. *([14], Proposition 1.4.7) Let S be an inverse semigroup.*

(1) The relation $\leqslant$ is a partial order on S.

(2) For idempotents $e, f \in S$, we have that $e \leqslant f$ if and only if $e = ef = fe$.

(3) If $s \leqslant t$ and $u \leqslant v$ then $su \leqslant tv$.

(4) If $s \leqslant t$ then $s^{-1}s \leqslant t^{-1}t$ and $ss^{-1} \leqslant tt^{-1}$.

(5) $E(S)$ is an order ideal of S.

Proof. We prove some parts, and refer to [14, Proposition 1.4.7], for the remaining details.
(1). The relation is reflexive $s \leqslant s$, since $s = s(s^{-1}s)$. Now let $s \leqslant t$, and $t \leqslant s$. Hence $s = t(s^{-1}s)$ and $t = s(t^{-1}t)$, so that

$$s = ts^{-1}s = st^{-1}ts^{-1}s = st^{-1}t = t.$$

Thus the relation is antisymmetric. Now suppose $s \leqslant t,$ and $t \leqslant v$. Then $s = te$ and $t = vf$ for some idempotents e and f. Therefore $s = te = (vf)e = v(fe)$. Hence $s \leqslant v$ and so the relation is transitive.

(2). Assume that $e \leqslant f$. Then $e = fh$ for some idempotent h. Thus $fe = f^2h = fh = e$, and so $e = fe = ef$ since idempotents commute. The converse is immediate.

(5). Let $s \leqslant t$ where $s, t \in S$ and $t \in E(S)$. From the definition of the natural partial order on S, we get $s = te$ where $e \in E(S)$. Then $te \in E(S)$, and so $s \in E(S)$. Hence $E(S)$ is an order ideal. $\square$

Let $(P, \leqslant)$ be a partially ordered set. If $z \leqslant x, y$ then z is said to be a *lower bound* of x and y. If z is the largest of the lower bounds, then it is called the *greatest lower bound* and denoted by $x \wedge y$. A *meet semilattice* is a poset in which every pair of elements has a greatest lower bound. The next result describes the natural partial order among the idempotents of inverse semigroups.

Proposition 1.3.3. *([14], Proposition 1.4.8) Let S be any semigroup. Define a relation $\leqslant$ on $E(S)$ by*

$$e \leqslant f \quad \textit{if and only if} \quad e = ef = fe\,.$$

Then $\leqslant$ is a partial order on $E(S)$. If S is an inverse semigroup, then $(E(S), \leqslant)$ is a meet semilattice.

Proof. We prove the second part: the proof of the first part can be found in [14, Proposition 1.4.8].

Assume that S is an inverse semigroup. Let $e, f \in E(S)$. Then

$$(ef)\,e = (fe)\,e = fe^2 = fe = ef,$$

by commutativity. Thus $ef \leqslant e$. Similarly, $ef \leqslant f$, since $(ef)f = ef^2 = ef$. Thus ef is a lower bound for e and f.

Now let $h \in E(S)$ be another lower bound for e and f; i.e. $h \leqslant e, f$. Since $h \leqslant e$, this implies that $h = he = eh$, and $h \leqslant f$, implies that $h = hf = fh$. So $h(ef) = (he)f = hf = h$. Thus $h \leqslant ef$, and so ef is the greatest lower bound of e and f. Therefore, $e \wedge f = ef$. It follows that $(E(S), \leqslant)$ is a meet semilattice. □

As a result of this proposition, the set of idempotents $E(S)$ of an inverse semigroup is called a *semilattice of idempotents*.

Proposition 1.3.4. *([14], Proposition 1.4.9) Meet semilattices are precisely the inverse semigroups in which every element is an idempotent.*

The following simple lemma will be of use later on.

Lemma 1.3.5. *Let S be an inverse semigroup. Then an inverse subsemigroup F of S is an order ideal in S if and only if $E(F)$ is an order ideal in $E(S)$.*

Proof. We start by assuming that $E(F)$ is an order ideal in $E(S)$. Take an element $t \in F$, such that $s \leqslant t,$ and we shall deduce that $s \in F$.

Now $s \leqslant t$ implies that $s^{-1} \leqslant t^{-1}$, by the characterization of the natural partial order. Then from Proposition 1.3.2(4), we have that $ss^{-1} \leqslant tt^{-1}$, and of course $tt^{-1} \in F$. Since $E(F)$ is an order ideal, $ss^{-1} \in E(F)$. By Theorem 1.3.1 we have $s = ss^{-1}t$ and so $s \in F$. Therefore F is an order ideal.

To verify the converse, suppose that F is an order ideal in S. Suppose that $f \in E(F)$ and that $s \leqslant f$. Then $s \in F$, but by Proposition 1.3.2(5) we also have $s \in E(S)$. Hence $s \in E(F)$ and $E(F)$ is an order ideal. □

Let S be an inverse semigroup, and let H be a subset of S. The *closure* $(H)^{\uparrow}$ of H in S, is defined by

$$(H)^{\uparrow} = \{s \in S : (\exists h \in H)\, h \leqslant s\} \ .$$

Therefore, the subset H will be called *closed (upwards)* if $(H)^{\uparrow} = H$.

Proposition 1.3.6. *([11], Proposition 5.2.2) If H is an inverse subsemigroup of an inverse semigroup S, then $(H)^{\uparrow}$ is a closed inverse subsemigroup of S.*

Lemma 1.3.7. *Suppose that $s = pq$ in an inverse semigroup S. Then $ss^{-1} \leqslant pp^{-1}$. Hence if s is an element of some closed inverse subsemigroup of S, then so is p.*

Proof. We have $ss^{-1} = pqq^{-1}p^{-1} \leqslant pp^{-1}$. □

1.4 Compatibility Relations

Let S be an inverse semigroup.

For all $s, t \in S$, the *left compatibility relation* is defined by

$$s \sim_l t \Leftrightarrow st^{-1} \in E(S),$$

the *right compatibility relation* is defined by

$$s \sim_r t \Leftrightarrow s^{-1}t \in E(S),$$

and the *compatibility relation,* the intersection of the above two relations, is defined by

$$s \sim t \Leftrightarrow st^{-1}, s^{-1}t \in E(S)\,.$$

It is quite clear that the above three relations are reflexive and symmetric, but none of them need be transitive. However, the characterization of the inverse semigroups having a transitive compatibility relation is given in Theorem 1.4.2.

Lemma 1.4.1. *([14], Lemma 1.4.11 and 1.4.12) Let S be an inverse semigroup and let $s, t \in S$.*

(1) $s \sim_l t$, if and only if the greatest lower bound $s \wedge t$ of s and t exists and $(s \wedge t)^{-1}\,(s \wedge t) = s^{-1}\,st^{-1}\,t$.

(2) $s \sim_r t$, if and only if the greatest lower bound $s \wedge t$ of s and t exists and $(s \wedge t)\,(s \wedge t)^{-1} = ss^{-1}\,tt^{-1}$.

(3) $s \sim t$, if and only if the greatest lower bound $s \wedge t$ of s and t exists and

$$(s \wedge t)^{-1}\,(s \wedge t) = s^{-1}st^{-1}t \;\; \textit{and} \;\; (s \wedge t)\,(s \wedge t)^{-1} = ss^{-1}\,tt^{-1}.$$

(4) If $s \sim_l t$, then

$$s \wedge t = st^{-1}t = ts^{-1}t = ts^{-1}s = st^{-1}s\,.$$

(5) If $s \sim_r t$, then

$$s \wedge t = ss^{-1}t = st^{-1}s = tt^{-1}s = ts^{-1}t\,.$$

(6) If $s \sim t$, then

$$s \wedge t = st^{-1}t = ts^{-1}t = ts^{-1}s = st^{-1}s = ss^{-1}t = tt^{-1}s\,.$$

Definition An inverse semigroup is *E–unitary* if, whenever e is an idempotent and $e \leqslant s$, then s is an idempotent. In other words, S is E–unitary if and only if $E(S)$ is a closed inverse subsemigroup of S.
An inverse semigroup S with a zero cannot be E–unitary unless $S = E(S)$. Instead we define S to be *E^*–unitary* if, whenever we have non-zero elements $s, t \in S$ with $s \leqslant t$ and $s \in E(S)$ then $t \in E(S)$.

Theorem 1.4.2. *([14], Lemma 2.4.4) Let S be an inverse semigroup. Then the compatibility relation is transitive if and only if S is $E-$unitary.*

1.5 Congruence

Definition An equivalence relation is called a *left congruence* if $(a, b) \in \rho$ implies that $(ca, cb) \in \rho$ for any $c \in S$. *Right congruences* are defined dually. So, an equivalence relation is a *congruence* if it is a left and right congruence. Also, other version of this definition is the following:
A *congruence* on a semigroup S is an equivalence relation ρ on S such that $(a, b), (c, d) \in \rho$ implies that $(ac, bd) \in \rho$. Therefore, every semigroup homomorphism determines a congruence on its domain. The reverse is also true. Let ρ be an arbitrary congruence on the semigroup S. We denote the set of $\rho-$equivalence classes (or congruence classes) by S/ρ. A binary operation may be defined on the set S/ρ by mapping $(\rho(a), \rho(b))$ to $\rho(ab)$.

This mapping is well-defined because ρ is a congruence.

Definition Let $\theta : S \to T$ be a homomorphism of semigroups. The *kernel* of θ is the relation $ker\,\theta$ defined on S by:

$$ker\,\theta = \{(a,b) \in S \times S : \theta(a) = \theta(b)\}\,.$$

Clearly, $ker\,\theta$ is an equivalence relation on S and has this additional property:

$$(a,b),(c,d) \in ker\,\theta \Longrightarrow (ac,bd) \in ker\,\theta\,.$$

Definition A homomorphism of inverse semigroups is said to be *idempotent pure* if the inverse images of idempotents consist only of idempotents.

Definition A congruence ρ on an inverse semigroup S is said to be *idempotent pure* if $a \in S$ and $e \in E(S)$ and $(a,e) \in \rho$ then a is an idempotent.

1.6 The Minimum Group Congruence

Definition The relation σ is defined on the inverse semigroup S by

$$s\,\sigma\,t \text{ if and only if there exists } u \leqslant s,t$$

for all $s,t \in S$. Then σ is a congruence on S and S/σ is a group.

Theorem 1.6.1. *Let S be an inverse semigroup.*

(1) σ is the smallest congruence on S containing the compatibility relation.

(2) S/σ is a group.

(3) If ρ is any congruence on S such that S/σ is a group, then $\sigma \subseteq \rho$.

The congruence σ is, in fact, the *minimum group congruence*. We shall denote the quotient S/σ by $\widehat{S}$.

Lemma 1.6.2. *Let F be a full inverse subsemigroup of the inverse semigroup S. Then the group homomorphism $\widehat{F} \to \widehat{S}$ induced by the inclusion $F \hookrightarrow S$ is injective.*

Proof. Denote the minimum group congruences on F and S by σ_F and σ_S respectively. Suppose that $a, b \in F$ and that $a\,\sigma_S\,b$. Then there exists $u \in S$ with $u \leqslant a$ and $u \leqslant b$. By Theorem 1.3.1(4) we have $u = uu^{-1}a = uu^{-1}b$. Since $uu^{-1} \in F$, we deduce that $u \in F$ and so $a\,\sigma_F\,b$. □

Theorem 1.6.3. *([14], Theorem 2.4.6) Let S be an inverse semigroup. Then the minimum group congruence is idempotent-pure if and only if S is $E-$unitary.*

1.7 Green's Relations

Green's relations were introduced by J. A. Green in [10] and have played a fundamental role in the development of semigroup theory. The characterization of Green's relations in arbitrary semigroups is given by the notion of *ideal* mentioned in Section 1.1. More precisely, the use of *principal ideals* provides a link with the usual way of defining Green's equivalences in general semigroups.

Definition For each element s of a semigroup S, the smallest left ideal of S containing s is $Ss \cup \{s\}$, which it is convenient to denote by S^1s. We shall call it the *principal left ideal generated by* s.

The following notations will be standard:

$$S^1s = Ss \cup \{s\}\,,$$

$$s\,S^1 = sS \cup \{s\}\,,$$

$$S^1sS^1 = SsS \cup Ss \cup sS \cup \{s\}\,.$$

Notice that S^1s, $s\,S^1$ and S^1sS^1 are all subsets; (indeed subsemigroups), of S (a semigroup), and they do not contain the element 1.

In inverse semigroups, we always have $s \in Ss$, and $s \in sS$, since $s = (ss^{-1})s = s(s^{-1}s)$. Therefore, the principal right ideal containing s is sS, the principal left ideal containing s is Ss, and the principal two-sided ideal containing s is SsS.

In groups, the situation is very different. Evidently, for any group element g, we have $Gg = G = gG$. Therefore, ideals do not have any role in group theory.

Definition An equivalence $\mathcal{L}$ on S (a semigroup) is defined by the rule that $s\,\mathcal{L}\,t$ if and only if s and t generate the same principal left ideal, that is, if and only if $S^1s = S^1t$. In a similar way, we define the equivalence $\mathcal{R}$ by the rule that $s\,\mathcal{R}\,t$ if and only if s and t generate the same principal right ideal, that is, if and only if $s\,S^1 = t\,S^1$.

Lemma 1.7.1. *([14], Section 3.2) In an inverse semigroup S, the relations $\mathcal{L}$ and $\mathcal{R}$ are*

$$(s,t) \in \mathcal{L} \Longleftrightarrow s^{-1}s = t^{-1}t \quad \textit{and} \quad (s,t) \in \mathcal{R} \Longleftrightarrow ss^{-1} = tt^{-1}\,.$$

Both $\mathcal{L}$ and $\mathcal{R}$ are equivalence relations, and the equivalence relation $\mathcal{H}$ is defined by $\mathcal{H} = \mathcal{L} \cap \mathcal{R}$. An inverse semigroup is called *combinatorial,* if $s^{-1}s = t^{-1}t$ and $ss^{-1} = tt^{-1}$ implies that $s = t$. Equivalently, combinatorial inverse semigroups can be defined as those in which Green's $\mathcal{H}-$relation is the equality relation.

Proposition 1.7.2. *([14], Proposition 3.2.4) Let S be any semigroup. Then*

$$\mathcal{D} = \mathcal{L} \circ \mathcal{R} = \mathcal{R} \circ \mathcal{L},$$

where $\mathcal{D}$ is an equivalence relation.

Proposition 1.7.3. *([14], Proposition) In an inverse semigroup S,*

$$s\,\mathcal{D}\,t \Longleftrightarrow \exists\, z \;\; \textit{such that} \;\; s^{-1}s = zz^{-1} \;\; \textit{and} \;\; z^{-1}z = tt^{-1}.$$

An inverse semigroup S is said to be *bisimple* if and only if Green's $\mathcal{D}-$relation has one class.

For inverse semigroup with zero, there is a modification of the definition of bisimplicity which is more appropriate. Let S be an inverse semigroup with zero. If $s\,\mathcal{L}\,0$ then $s^{-1}s = 0$ and so $s = s(s^{-1}s) = s0 = 0$. Similarly, If $s\,\mathcal{R}\,0$ then $s = 0$. Therefore, we can conclude that the zero forms a $\mathcal{D}-$class on its own. An inverse semigroup with zero S is said to be *0-bisimple*, if it has exactly two $\mathcal{D}-$classes. Moreover, a *0-bisimple* inverse semigroup with zero S can be defined as follows: if for any two non-zero idempotents e and f, there exists an element s such that $e = ss^{-1}$ and $f = s^{-1}s$.
Although bisimple inverse semigroups have only one $\mathcal{D}-$class, they do not generally have an elementary structure. This follows from the result proved by Reilly [22], which says that every inverse semigroup could be embedded in a bisimple inverse monoid.

The final Green's relation $\mathcal{J}$, is defined with the aid of principal two-sided ideals of S. The principal two-sided ideal of S generated by s is $S^1 s\, S^1$, and then we can define the equivalence $\mathcal{J}$ by the rule that

$$s\,\mathcal{J}\,t \Longleftrightarrow S^1sS^1 = S^1tS^1,$$

that is to say, if and only if there are $x, y, u, v \in S^1$, such that

$$xsy = t\,, \quad utv = s.$$

It is clear that $\mathcal{L} \subseteq \mathcal{J}$ and $\mathcal{R} \subseteq \mathcal{J}$. Since $\mathcal{D}$ is the smallest equivalence contained in $\mathcal{L}$ and $\mathcal{R}$, we have $\mathcal{D} \subseteq \mathcal{J}$. In groups, we have equality $\mathcal{H} = \mathcal{L} = \mathcal{R} = \mathcal{D} = \mathcal{J} = G \times G$.

The next result is particularly important in the theory of inverse semigroups.

Theorem 1.7.4. *([11], Theorem 5.1.1) Let S be a semigroup. Then the following statements are equivalent:*

(1) S is an inverse semigroup;

(2) S is regular, and its idempotents commute;

(3) every $\mathcal{L}-$class and every $\mathcal{R}-$class contains exactly one idempotent;

(4) *every element of S has a unique inverse.*

The next proposition describes the relationship between the natural partial order and the three Green's relations $\mathcal{L}$, $\mathcal{R}$ and $\mathcal{H}$.

Proposition 1.7.5. *Let S be an inverse semigroup and let $s, t \in S$.*

(1) *$s \,\mathcal{L}\, t$ and $s \leqslant t$ implies that $s = t$.*

(2) *$s \,\mathcal{R}\, t$ and $s \leqslant t$ implies that $s = t$.*

(3) *$s \,\mathcal{H}\, t$ and $s \leqslant t$ implies that $s = t$.*

1.8 Clifford Semigroups

One of the natural examples of semigroups are Clifford semigroups. This section concentrates on the basic structure of these semigroups and obtains some information, see [14, Section 5.2].

Definition For every inverse semigroup S, we define the set $Z(E(S))$, the *centralizer of the idempotents,* to consist of all elements of S which commute with every idempotent. If $Z(E(S)) = S$ the semigroup is said to be *Clifford,* in other words, it is a semigroup with central idempotents. In fact, $Z(E(S))$ is always a Clifford semigroup, possibly consisting only of idempotents.

Now, we will obtain some information about the structure of these semigroups. The basic construction of Clifford semigroups is the following:

Let $(E, \leqslant)$ be a meet semilattice, and let $\{G_e : e \in E\}$ be a family of disjoint groups indexed by the elements of E, the identity of G_e being denoted by 1_e. For each pair e, f of elements of E where $e \geqslant f$ let $\phi_{e,f} : G_e \to G_f$ be a group homomorphism, such that the following two axioms hold:

(PG1) $\phi_{e,e}$ is the identity homomorphism on G_e.

(PG2) if $e \geqslant f \geqslant g$ then $\phi_{f,g}\,\phi_{e,f} = \phi_{e,g}$.

We call such a family

$$(G_e, \phi_{e,f}) = (\{G_e : \, e \in E\}, \{\phi_{e,f} : \, e, f \in E, \, f \leqslant e\})$$

a presheaf of groups (over the semilattice E).

Proposition 1.8.1. *([14], Proposition 5.2.11) Let $(G_e, \phi_{e,f})$ be a presheaf of groups. Let $S = S(G_e, \phi_{e,f})$ be the union of the G_e equipped with the product $\otimes$ on S defined by:*

$$x \otimes y = \phi_{e,e\wedge f}(x)\;\phi_{f,e\wedge f}(y)\,,$$

where $x \in G_e$ and $y \in G_f$. Then $(S, \otimes)$ is a Clifford semigroup.

The semigroup S constructed in Proposition 1.8.1 is called a *strong semilattice of groups.* In the next result, we show that every Clifford semigroup is isomorphic to a strong semilattice of groups.

Theorem 1.8.2. *([14], Theorem 5.2.12) Let S be an inverse semigroup. Then the following are equivalent:*

(1) S is a Clifford semigroup.

(2) For every $s \in S$, we have that $s^{-1}s = ss^{-1}$.

(3) Every $\mathcal{H}-$class is a group.

(4) Every $\mu-$class is a group.

(5) S is isomorphic to a strong semilattice of groups.

The following proposition presents a natural class of inverse semigroups that are automatically Clifford.

Proposition 1.8.3. *([14], Proposition 5.2.13) Let S be an inverse semigroup whose idempotents form a finite chain. Then S is a Clifford semigroup.*

The semigroups with finite chains of idempotents are often called *finite chain of groups*.

1.9 Symmetric Inverse Monoid

We have so far encountered groups, semilattices and Clifford semigroups as examples of inverse semigroups. We find a more representative example in the following:

Example 1.9.1. An enlightening example of an inverse semigroup is the *symmetric inverse monoid* on a set X, denoted by $\mathscr{I}(X)$. The *monoid* $\mathscr{I}(X)$ consists of all partial bijections from subsets of X to subsets of X with respect to composition of partial bijections. In other words, consider $\mathscr{I}(X)$ to be the set of all bijections $\alpha : A \longrightarrow B$, where $A, B \subseteq X$. The *multiplication* of partial bijections is as follows:
if α and β are two partial one-one maps, then $x\alpha\beta = (x\alpha)\beta$ whenever this makes sense, that is if $x \in \mathrm{dom}(\alpha)$ and $x\alpha \in \mathrm{dom}(\beta)$.
The empty subset corresponds to the zero of $\mathscr{I}(X)$, where a product $\alpha\beta$ of two partial bijections on X, is zero, if $\mathrm{ran}(\alpha) \cap \mathrm{dom}(\beta) = \emptyset$.

Proposition 1.9.2. *The idempotents of $\mathscr{I}(X)$ are precisely the partial identity maps on subsets of X, and if $|X| = n$, then there are 2^n of them. In particular, the idempotents form a commutative subsemigroup.*

Proposition 1.9.3. *The composition in $\mathscr{I}(X)$ is associative; that is $\alpha(\beta\gamma) = (\alpha\beta)\gamma$ for all $\alpha, \beta, \gamma \in \mathscr{I}(X)$. Moreover, for every element α there exists a unique element $\beta = \alpha^{-1}$ satisfying $\alpha\beta\alpha = \alpha$ and $\beta\alpha\beta = \beta$, and in which the idempotents commute with one another. Therefore, $\mathscr{I}(X)$ defines an inverse semigroup. In addition, $\mathscr{I}(X)$ has an identity element which makes $\mathscr{I}(X)$ an inverse monoid.*

This follows from the fact that the composition of mappings is always associative. However, this composition is not commutative in general, since $\alpha\beta \neq \beta\alpha$.

Proposition 1.9.4. *Let $\mathscr{I}(X)$ be the symmetric inverse monoid.*

(1) If $\alpha : A \longrightarrow B$ is a partial bijection in $\mathscr{I}(X)$, then there exists an inverse partial

bijection $\alpha^{-1} : B \longrightarrow A$, *where* $A, B \subseteq X$ *such that* $\alpha\alpha^{-1} = 1_A$ *(partial identity on* A*) and* $\alpha^{-1}\alpha = 1_B$ *(partial identity on* B*).*

(2) For the partial identity on A, 1_A *determines* A *and so* $1_A = 1_B$ *requires that* $A = B$.

(3) The idempotents in $\mathscr{I}(X)$ *are partial identities and* $1_A\, 1_B = 1_{A\cap B} = 1_B\, 1_A$.

Theorem 1.9.5. *The natural partial order on* $\mathscr{I}(X)$ *is the restriction of the domain of a partial bijection, that is*

$$\alpha \leqslant \beta \quad \textit{if and only if} \quad \textit{dom}\,(\alpha) \subseteq \textit{dom}\,(\beta) \;\; \textit{and} \;\; x\,\alpha = x\,\beta \;\; \textit{for all} \;\; x \in \textit{dom}\,(\alpha).$$

Theorem 1.9.6. *([11], Chapter 5, Ex. 3) If* $|X| = n$ *then the number of elements in* $\mathscr{I}(X)$ *is given by*

$$|\mathscr{I}(X)| = \sum_{r=0}^{n} \binom{n}{r}^2 r!\,.$$

In a real sense, $\mathscr{I}(X)$ is the appropriate analogue in inverse semigroup theory of the symmetric group in group theory and the full transformation semigroup in semigroup theory. Just as every group can be embedded up to isomorphism in a symmetric group (Cayley's Theorem) and every semigroup can be embedded up to isomorphism in a full transformation semigroup [11, Theorem 1.1.2], so every inverse semigroup can be embedded in a symmetric inverse semigroup. The inverse semigroup result is proved in a way that is similar in spirit to the earlier proofs, but is a little more complicated. We state the analogue of Cayley's Theorem, a result due to Wagner [31] (1952) and (independently) to Preston [21] (1954c):

Theorem 1.9.7. *([11], Theorem 5.1.7) Let* S *be an inverse semigroup. Then there exists a symmetric inverse semigroup* $\mathscr{I}(X)$ *and a monomorphism* ϕ *from* S *into* $\mathscr{I}(X)$*.*

1.10 Inverse Semigroups and Inductive Groupoids

A *groupoid* G is a small category in which every morphism is invertible. We consider a groupoid as an algebraic structure following [14, Chapter 4]: the elements are the morphisms, and composition is an associative partial binary operation. The set of identities in G is denoted $\mathrm{Ob}(G)$, and an element $g \in G$ has domain $\mathrm{dom}\, g = gg^{-1}$ and range $\mathrm{ran}\, g = g^{-1}g$. (Note that this follows the conventions of [14, Chapter 4]). For each $x \in \mathrm{Ob}(G)$ the set $G(x) = \{g \in G : \mathrm{dom}\, g = x = \mathrm{ran}\, g\}$ is a subgroup of G, called the *local subgroup* at x. A groupoid G is *connected* if, for any $x, y \in \mathrm{Ob}(G)$, there exists $g \in G$ with $\mathrm{dom}\, g = x$ and $\mathrm{ran}\, g = y$. In a connected groupoid G, all local subgroups are isomorphic, and for any such local subgroup L there is an isomorphism $G \cong \mathrm{Ob}(G) \times L \times \mathrm{Ob}(G)$, where the latter set carries the groupoid composition $(x, k, y)\,(y, l, z) = (x, kl, z)$.

An *ordered groupoid* $(G, \leqslant)$ is a groupoid G with a partial order $\leqslant$ satisfying the following axioms:

(OG1) for all $g, h \in G$, if $g \leqslant h$ then $g^{-1} \leqslant h^{-1}$,

(OG2) if $g_1 \leqslant g_2$, $h_1 \leqslant h_2$ and if the compositions g_1h_1 and g_2h_2 are defined, then $g_1h_1 \leqslant g_2h_2$,

(OG3) if $g \in G$ and x is an identity of G with $x \leqslant \mathrm{dom}\, g$, there exists a unique element $(x|g)$, called the *restriction* of g to x, such that $\mathrm{dom}(x|g) = x$ and $(x|g) \leqslant g$,

As a consequence of (OG3) we also have:

(OG3*) if $g \in G$ and y is an identity of G with $y \leqslant \mathrm{ran}\, g$, there exists a unique element $(g|y)$, called the *corestriction* of g to y, such that $\mathrm{ran}(g|y) = y$ and $(g|y) \leqslant g$,

since the corestriction of g to y may be defined as $(y|g^{-1})^{-1}$.

Let G be an ordered groupoid and let $a, b \in G$. If $\operatorname{ran} a$ and $\operatorname{dom} b$ have a greatest lower bound $\ell \in \operatorname{Ob}(G)$, then we may define the *pseudoproduct* of a and b in G as:

$$a \otimes b = (a|\ell)\,(\ell|b),$$

where the right-hand side is now a composition defined in G. As Lawson shows in [14, Lemma 4.1.6], this is a partially defined associative operation on G.

If $\operatorname{Ob}(G)$ is a meet semilattice then G is called an *inductive* groupoid. The pseudoproduct is then everywhere defined and $(G, \otimes)$ is an inverse semigroup. On the other hand, given an inverse semigroup S with semilattice of idempotents $E(S)$, then S is a poset under the natural partial order, and the restriction of its multiplication to the partial composition

$$a \cdot b = ab \in S \;\text{ defined when }\; a^{-1}a = bb^{-1}$$

gives S the structure of an inductive groupoid, which we denote by $\vec{S}$, with $\operatorname{Ob}(\vec{S}) = E(S)$. These constructions give rise to an isomorphism between the categories of inverse semigroups and inductive groupoids: this is the *Ehresmann-Schein-Nambooripad Theorem* [14, Theorem 4.1.8].

It is sometimes useful to adopt a less formal version of this correspondence, and to think of an element s of an inverse semigroup S as an arrow joining the idempotent ss^{-1} to the idempotent $s^{-1}s$.

1.11 Free Monoids and Free Inverse Monoids

Let A be a set, and let A^* be the set of all finite strings of elements of A. The empty string, denoted by ε, is an element of A^*. The length of a string $w \in A^*$ is denoted by $|w|$. Then A^* is a monoid under the operation of concatenation of strings, with identity element ε, and is called the *free monoid* on A. Note that the strings of length 1 are the elements of A, so that $A \subset A^*$. The freeness property possessed by A^* is the extension of functions to monoid homomorphisms: given any function $f : A \to M$ from A to a monoid M, there exists a unique monoid homomorphism $f^* : A^* \to M$ such that, for all $a \in A,\ af^* = af$.

Let F be a group, let X be a set, and let $\theta : X \to F$ be a mapping. Then F is called a *free group* on X (with respect to θ); denoted by $F(X)$, if for every group G and every mapping $f : X \to G$ there exists a unique homomorphism $\phi : F \to G$ such that $\phi\theta = f$. Another customary way of abbreviating this definition is to say: every mapping $X \to G$ can be uniquely extended to a homomorphism $F \to G$.

The *free inverse monoid* $\mathrm{FIM}(X)$ on a set X satisfies the following freeness property: given any function $f : X \to M$ from X to an inverse monoid M, there exists a unique inverse monoid homomorphism $f^{\#} : \mathrm{FIM}(X) \to M$ such that, for all $x \in X,\ xf^{\#} = xf$. The elements of $\mathrm{FIM}(X)$ were described by Munn [19] using what are now called ***Munn trees***: see [14, Section 6.4].

For each word $w \in (X \cup X^{-1})^*,$ we let $MT(w)$ be the finite subtree of the Cayley graph $\Gamma(X)$ of the free group $F(X)$ traversed by reading the path in $\Gamma(X)$ labelled by $w,$ starting at the vertex 1 and ending at $r(w)$ (the reduced form of w). Munn's Theorem [19, Theorem 1.4] then states that, for all words $u, v \in (X \cup X^{-1})^*$, $u\rho v$ (i.e. $u = v$ in $\mathrm{FIM}(X)$) if and only if $(MT(u), r(u)) = (MT(v), r(v)).$ The tree $MT(w)$ is referred to as the ***Munn tree*** of w.

Let $F(X)$ be the free group on X. The Cayley graph Γ of $F(X)$ is a tree with vertex set $V(\Gamma) = F(X)$, and an element of $\mathrm{FIM}(X)$ is a pair (P, u) where P is a finite connected subtree of Γ with $1 \in V(P)$, and $u \in V(P)$. The multiplication in $\mathrm{FIM}(X)$ is then given by:

$$(P, u)(Q, v) = (P \cup uQ, uv).$$

We note that a generator $x \in X$ is then represented by the pair (e_x, x) where e_x is the directed edge in Γ from 1 to x. The natural partial order on $\mathrm{FIM}(X)$ is given by

$$(P, u) \leqslant (Q, v) \iff P \supseteq Q \text{ and } u = v\,.$$

Chapter 2

Cosets in Inverse Semigroups

The main goal of this chapter is to state and prove the Index Formula for inverse semigroups, Theorem 2.3.1 and the inverse semigroup analogue of Hall's Theorem, Theorem 2.5.2.

Areas of group and semigroup theory necessary for the remainder of this chapter are briefly covered in the next section, which also makes definitions and establishes notations. For further background information on the general theory of cosets in groups and inverse semigroups and the actions of groups and inverse semigroups, see [24, Chapter 3,4], [14, Chapter 1] and [16].

2.1 Group Cosets and Actions

Definition Let G be a group, and $H \leqslant G$ be a subgroup. Then the *right cosets* Hg of H in G, are the subsets $Hg = \{hg : h \in H\}$ with $g \in G$ (where each g determines a subset Hg). Similarly, the *left cosets* gH of H in G are the subsets $gH = \{gh : h \in H\}$ with $g \in G$.

Theorem 2.1.1. *Let G be a group, and $H \leqslant G$ be a subgroup. Then $Hg_1 = Hg_2$ if and only if $g_1 g_2^{-1} \in H$, and $g_1 H = g_2 H$ if and only if $g_1^{-1} g_2 \in H$.*

Definition Let G be a group, and X be a set. A homomorphism

$$\theta : G \longrightarrow \Sigma_X$$

into the symmetric group (group of permutations) on X is called a *permutation representation*. We call such a representation *faithful* if θ is injective.

Definition A (right) *group action* of the group G on the set X is determined by a function $X \times G \longrightarrow X$, where we write $(x, g) \mapsto x \lhd g,$ satisfying two axioms:

- $x \lhd gh = (x \lhd g) \lhd h$
- $x \lhd 1 = x \qquad (x \in X;\ g, h \in G).$

We shall also say that G *acts* on the set X and that X is a (right) G-*space*. Left G-spaces can be defined dually.

Definition Let G act on the set X, and let $x \in X$. Set $\mathrm{stab}(x) = \{g \in G : x \lhd g = x\}$. Then $\mathrm{stab}(x)$ is a subgroup of G, called the *stabilizer* of x in G.

Definition Let G be a group, and $H \leqslant G$ be a subgroup. Then a complete set of representatives of the right cosets Hg of H in G is a set R consisting of one element from each coset. The element in R coming from the coset Hg is denoted by $\overline{g},$ and is called the *representative* of Hg. In addition, if $1 \in R$, then R is called a *right transversal* of H in G.

Example 2.1.2. (Actions on Cosets) Let H be a subgroup of a group G. Then there is a natural action of G on the set of right cosets G/H by right multiplication: $(Ha) \lhd g = H(ag)$.

Theorem 2.1.3. *There is a bijective correspondence between actions of a group G on a set X and representations of G by means of permutations on X.*

2.2 Closed Inverse Subsemigroups and Cosets

First of all, it is necessary to assemble various definitions and results to establish the background information for this chapter. In this section, we define cosets for closed inverse subsemigroups following Schein's convention, see [26], the action of inverse semigroups on sets, and then establish the principal characterization of cosets in this convention. Some of the definitions are extracted from [16].

Definition Let S be an inverse semigroup. For $a, b, c \in S$ we define $\langle a, b, c \rangle = ab^{-1}c$. We call this ternary operation the *heap* operation on S, see [4]. An *atlas* in S is a subset $A \subseteq S$ such that $AA^{-1}A \subseteq A$: that is, A is closed under the heap operation. Since, for all $a \in A$ we have $\langle a, a, a \rangle = a$, we see that A is an atlas if and only if $AA^{-1}A = A$. For a more detailed introduction to the concept of atlas, and the connection with cosets to be explained below, we refer to [13].

Amongst the basic properties of atlases we have:

Lemma 2.2.1. *Let S be an inverse semigroup.*

- *If A is an atlas in S then $A^{-1}A$ and AA^{-1} are inverse subsemigroups of S,*
- *If A and B are atlases in S then so is their intersection $A \cap B$,*
- *Every principal ideal in S is an atlas,*
- *Every local subsemigroup in S is an atlas.*

For groups, the concept of an atlas is the same as that of a coset.

Theorem 2.2.2. *([6], Theorem 1) $S \subset G$ is closed under the* heap *ternary operation $\langle a, b, c \rangle = ab^{-1}c$ if and only if S is a coset of some subgroup of G; indeed a right (left) coset of $S^{-1}S$ (SS^{-1}).*

Proof. For the first part, assume that S is a coset of $S^{-1}S$ (or SS^{-1}). Then $S = SS^{-1}s$ and so $ss^{-1}s \in S$. Also, if $S = s\,S^{-1}S$ then $ss^{-1}s \in S$. Hence S is closed.

For the second part, observe that if $S = s\,T(s \in S)$, then $T = s^{-1}S$ $(s \in S)$ and hence $T = S^{-1}S$. Similarly, $S = T\,s(s \in S)$, implies $T = S\,S^{-1}$. $\square$

Following the idea in Theorem 2.2.2, Schein [26] uses the heap operation to define cosets for inverse semigroups.

Definition A *coset* C in S is a closed atlas: that is, C is both upwards closed in the natural partial order on S and is closed under the heap operation $\langle \cdot \cdot \cdot \rangle$.

Definition A *representation* of an inverse semigroup S by means of partial bijections is a homomorphism $\rho : S \to \mathscr{I}(X)$ to the symmetric inverse monoid on a set X. If S is a monoid, we assume that the homomorphism is a monoid homomorphism. In this sense, we can define a corresponding notion of an action of the inverse semigroup S on the set X: the associated action is defined by $x \lhd s = x(s\rho)$. We shall also say that S *acts* on the set X and that X is a (right) S-*space*. Left S-*space* can be defined dually.
We shall use the words action and representation interchangeably, so the action of the inverse semigroup S on the set X implies the existence of an appropriate homomorphism from S to $\mathscr{I}(X)$.

Definition Let X be a set and $\mathscr{I}(X)$ its symmetric inverse monoid. Let $\rho : S \to \mathscr{I}(X)$ be a faithful representation of S on X, and write $x(s\rho)$ as $x \lhd s$. A *stationary subset* of S is a subset defined as follows. Given $x, y \in X$ we define

$$\mathrm{stat}(x, y) = \{s \in S : x \lhd s = y\}.$$

The principal characterizations of cosets that we need are due to Schein:

Theorem 2.2.3. *([26], Theorem 3) Let C be a non-empty subset of an inverse semigroup S. Then the following are equivalent:*

(1) C is a coset,

(2) there exists a closed inverse subsemigroup F of S such that, for all $s \in C$, we have $ss^{-1} \in F$ and $C = (Fs)^{\uparrow}$.

(3) there exists a closed inverse subsemigroup K of S such that, for all $s \in C$, we have $s^{-1}s \in K$ and $C = (sK)^{\uparrow}$.

Proof. ($1 \Longrightarrow 2$) Let $Q = CC^{-1} = \{ab^{-1} : a, b \in C\}$. Then Q is an inverse subsemigroup of S, since, for all $a, b, c, d \in C$ we have

- $(ab^{-1})(cd^{-1}) = (ab^{-1}c)d^{-1} = \langle a, b, c\rangle d^{-1} \in Q$,
- $(ab^{-1})^{-1} = ba^{-1} \in Q$.

Set $F = (Q)^{\uparrow}$: then F is a closed inverse subsemigroup of S. Let $s \in C$. Obviously $ss^{-1} \in Q \subseteq F$. Moreover, given any $c \in C$ we have $c \geqslant c(s^{-1}s) = (cs^{-1})s \in Qs \subseteq Fs$, so that $C \subseteq (Fs)^{\uparrow}$. Conversely, if $x \in (Fs)^{\uparrow}$, we have $x \geqslant us$ for some $u \in F$, with $u \geqslant ab^{-1}$ for some $a, b \in C$. Hence $x \geqslant us \geqslant ab^{-1}c = \langle a, b, c\rangle \in C$. Since C is closed, $x \in C$ and therefore $(Fs)^{\uparrow} \subseteq C$.

($2 \Longrightarrow 1$) The subset $(Fs)^{\uparrow}$ is a coset, since it is closed by definition, and if $h_i \in (Fs)^{\uparrow}$ we have $h_i \geqslant t_i s$ for some $t_i \in F$. Then $\langle h_1, h_2, h_3\rangle = h_1 h_2^{-1} h_3 \geqslant t_1 ss^{-1} t_2^{-1} t_3 s \in Fs$ since $ss^{-1} \in F$. It follows that $(Fs)^{\uparrow}$ is closed under the heap operation $\langle \cdots \rangle$.

For ($1 \Longleftrightarrow 3$) we proceed in the same way, with $K = (C^{-1}C)^{\uparrow}$. $\square$

Proposition 2.2.4. *([26], Proposition 5) A coset C that contains an idempotent $e \in E(S)$ is an inverse subsemigroup of S: indeed, in this case $C = (CC^{-1})^{\uparrow} = F$.*

Proof. If $a, b \in C$ then $ab \geqslant aeb = \langle a, e, b\rangle \in C$ and since C is closed, we have $ab \in C$. Furthermore, $a^{-1} \geqslant ea^{-1}e = \langle e, a, e\rangle \in C$ and so $a^{-1} \in C$. Hence C is an inverse subsemigroup.

Now $ab^{-1} \in CC^{-1}$ and $ab^{-1} \geqslant ab^{-1}e = \langle a,b,e\rangle \in C$. Since C is closed we have $F = (CC^{-1})^{\uparrow} \subseteq C$. But if $x \in C$ then $x \geqslant xe \in CC^{-1}$ and so $x \in (CC^{-1})^{\uparrow} = F$. Therefore $C = (CC^{-1})^{\uparrow} = F$. □

Now if L is a closed inverse subsemigroup of S, a (right) *coset* of L is a subset of the form $(Ls)^{\uparrow}$ where $ss^{-1} \in L$. Suppose that C is such a coset: then Theorem 2.2.3 associates to C the closed inverse subsemigroup $(CC^{-1})^{\uparrow}$.

Proposition 2.2.5. *([26], Proposition 6) Let L be a closed inverse subsemigroup of S.*

(1) Suppose that C is a coset of L. Then $(CC^{-1})^{\uparrow} = L$.

(2) If $t \in C$ then $tt^{-1} \in L$ and $C = (Lt)^{\uparrow}$. Hence two cosets of L are either disjoint or they coincide.

(3) Two elements $a, b \in S$ belong to the same coset C of L if and only if $ab^{-1} \in L$.

(4) If $e \in E(S)$ and $e \in C$ then $C = L$.

Proof. (1) If $c_i \in C$ $(i = 1, 2)$ then there exists $l_i \in L$ such that $c_i \geqslant l_i s$. Hence $c_1c_2^{-1} \geqslant l_1ss^{-1}l_2^{-1} \in L$, and so $CC^{-1} \subseteq L$. Since L is closed, $(CC^{-1})^{\uparrow} \subseteq L$. On the other hand, for any $l \in L$ we have $l = ll^{-1}l \geqslant lss^{-1}l^{-1}l = (ls)(l^{-1}ls)^{-1} \in CC^{-1}$ and so $L \subseteq (CC^{-1})^{\uparrow}$.

(2) If $C = (Ls)^{\uparrow}$ and $t \in C$ then, for some $l \in L$ we have $t \geqslant ls$. Then $tt^{-1} \geqslant lss^{-1}l^{-1} \in L$, and since L is closed, $tt^{-1} \in L$. Moreover, if $u \in (Lt)^{\uparrow}$ then for some $k \in L$ we have $u \geqslant kt \geqslant kls$ and so $u \in (Ls)^{\uparrow}$. Hence if $t \in (Ls)^{\uparrow}$ then $(Lt)^{\uparrow} \subseteq (Ls)^{\uparrow}$. Now $ls = (ls)(ls)^{-1}t = lss^{-1}l^{-1}t$ and so $l^{-1}ls = l^{-1}lss^{-1}l^{-1}t = ss^{-1}l^{-1}t \in Lt$. Since $s \geqslant l^{-1}ls$, we deduce that $s \in (Lt)^{\uparrow}$. Hence $(Ls)^{\uparrow} \subseteq (Lt)^{\uparrow}$.

(3) Suppose that $a, b \in (Ls)^{\uparrow}$. Then for some $k, l \in L$ we have $a \geqslant ks$ and $b \geqslant ls$: hence $ab^{-1} \geqslant kss^{-1}l^{-1} \in L$ and so $ab^{-1} \in L$. On the other hand, suppose that $ab^{-1} \in L$. Then $aa^{-1} \geqslant a(b^{-1}b)a^{-1} = (ab^{-1})(ab^{-1})^{-1} \in L$, and similarly $bb^{-1} \in L$. We note that $a = (aa^{-1})a \in La$ and similarly $b \in Lb$. Then $a \geqslant a(b^{-1}b) = (ab^{-1})b$

and so $a \in (Lb)^{\uparrow}$. As in part (2), we deduce that $(La)^{\uparrow} \subset (Lb)^{\uparrow}$. By symmetry $(La)^{\uparrow} = (Lb)^{\uparrow}$ and this coset contains a and b.

(4) By applying the result of part (2) of this proposition, which tells us that: if $t \in (Ls)^{\uparrow}$, then $(Lt)^{\uparrow} = (Ls)^{\uparrow}$. Therefore, if $e \in (C)^{\uparrow}$ and $C = (Lt)^{\uparrow}$ then $(Le)^{\uparrow} = (Lt)^{\uparrow}$. But $(Le)^{\uparrow} = L$, and so $(Lt)^{\uparrow} = L$. Hence, $C = L$. We obtain that the only coset containing an idempotent is the inverse subsemigroup L.

$\square$

Remark 2.2.6. Let F be a closed inverse subsemigroup of an inverse semigroup S. Then the union $\mathcal{U}$, of all the (right) cosets of F is a subset of S but need not be all of S. Moreover, the union of the cosets is not always a subsemigroup of S.

We will provide some fairly simple examples here which illustrate this remark.

Example 2.2.7. Fix a set X, and define the *Brandt semigroup* B_X as follows. As a set, we have

$$B_X = \{(x, y) : x, y \in X\} \cup \{0\}$$

and

$$(u, v)(x, y) = \begin{cases} (u, y) & \text{if } v = x \\ 0 & \text{if } v \neq x \end{cases}$$

and $0(x, y) = 0 = (x, y)0$. The idempotents of B_X are the elements (x, x) for $x \in X$ and 0. Hence $0 \leqslant (x, y)$ for all $x, y \in X$ and $(u, v) \leqslant (x, y)$ if and only if $(u, v) = (x, y)$. If a closed inverse semigroup F of B_X contains (x, y) with $x \neq y$ then $(x, y)(x, y) = 0 \in F$ and so $F = B_X$. Therefore the only proper closed inverse subsemigroups are the subsemigroups $E_x = \{(x, x)\}$ for $x \in X$. An element $(x, y) \in B_X$ then determines the coset

$$(E_x(x, y))^{\uparrow} = \{(x, y)\}.$$

Hence there are $|X|$ distinct cosets of E_x and their union is

$$\mathcal{U} = \{(x, y) : y \in X\}.$$

Example 2.2.8. Consider $\mathscr{I}(\{1,2\})$ which has 7 elements, see Section 1.9:

$$\mathscr{I}(\{1,2\}) = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 1 & * \end{pmatrix} \begin{pmatrix} 1 & 2 \\ * & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ * & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & * \end{pmatrix} \begin{pmatrix} 1 & 2 \\ * & * \end{pmatrix} \right\}.$$

Take $F = stab(1) = \{\alpha \in \mathscr{I}(X) :\ 1\alpha = 1\}$.

Therefore,

$$F = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 1 & * \end{pmatrix} \right\},$$

and by Theorem 1.9.5, F is a closed inverse subsemigroup. There are two idempotents in F, so if $\sigma\sigma^{-1} \in F$, then the domain of σ is either $\{1,2\}$ or $\{1\}$. Define

$$\sigma_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 2 \\ 1 & * \end{pmatrix}, \quad \sigma_4 = \begin{pmatrix} 1 & 2 \\ 2 & * \end{pmatrix}.$$

Hence,

$$(F\sigma_1)^{\uparrow} = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 1 & * \end{pmatrix} \right\} = F.$$

Similarly,

$$(F\sigma_2)^{\uparrow} = \left\{ \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & * \end{pmatrix} \right\},$$

whereas

$$(F\sigma_3)^{\uparrow} = F \quad and \quad (F\sigma_4)^{\uparrow} = (F\sigma_2)^{\uparrow}.$$

Hence, the cosets of F in $\mathscr{I}(\{1,2\})$ are: $(F\sigma_1)^{\uparrow}$ and $(F\sigma_2)^{\uparrow}$.

Furthermore, it can be seen in this example that there are still some elements of $\mathscr{I}(\{1,2\})$ which are not in any cosets of F, for instance :

$$\begin{pmatrix} 1 & 2 \\ * & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ * & 1 \end{pmatrix}.$$

Now we show that the union $\mathcal{U}$ of the cosets of F is not a subsemigroup of $\mathscr{I}(\{1,2\})$.

Let

$$\sigma = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \in (F\sigma_2)^{\uparrow} \quad and \quad \tau = \begin{pmatrix} 1 & 2 \\ 1 & * \end{pmatrix} \in (F\sigma_1)^{\uparrow}.$$

Therefore, it is evident that $\sigma\tau = \begin{pmatrix} 1 & 2 \\ * & 1 \end{pmatrix} \notin \mathcal{U}$. This yields the conclusion that the union of the cosets of F is not a subsemigroup of $\mathscr{I}(\{1,2\})$, since it is not closed under multiplication. Hence

$$(F\sigma_1)^{\uparrow} \cup (F\sigma_2)^{\uparrow} = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 1 & * \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & * \end{pmatrix} \right\}$$

is a proper subset of $\mathscr{I}(\{1,2\})$ but not a subsemigroup.

We shall now study the union of the cosets and see when it is equal to the whole inverse semigroup.

Theorem 2.2.9. *Let F be a closed inverse subsemigroup of an inverse semigroup S and let $\mathcal{U}$ be the union of all the (right) cosets of F in S. Then*

$$\mathcal{U} = \{ s \in S : s s^{-1} \in F \}$$

and therefore $\mathcal{U} = S$ if and only if F is full.

Proof. Since a coset $(Fu)^{\uparrow}$ exists if and only if $uu^{-1} \in F$ we have

$$\mathcal{U} = \bigcup_{\substack{u \in S \\ uu^{-1} \in F}} (Fu)^{\uparrow}.$$

Let $\mathcal{V} = \{s \in S : s s^{-1} \in F\}$. We claim that $\mathcal{U} = \mathcal{V}$.

If $v \in \mathcal{V}$ then $v \in (Fv)^{\uparrow}$ and so $\mathcal{V} \subseteq \mathcal{U}$. On the other hand, if $w \in \mathcal{U}$ then for some $v \in \mathcal{V}$ and $x \in F$ we have $w \geqslant xv$. Therefore $ww^{-1} \geqslant xvv^{-1}x^{-1} \in F$, since $x, x^{-1}, vv^{-1} \in F$, and since F is closed we deduce that $ww^{-1} \in F$, and hence that $w \in V$. Therefore $\mathcal{U} \subseteq \mathcal{V}$.

It is clear that $\mathcal{V} = S$ if and only if F is full. $\square$

Theorem 2.2.10. *The union of all the (right) cosets $\mathcal{U}$ of F is a closed inverse subsemigroup of S if and only if whenever $e \in E(F)$ and $s \in \mathcal{U}$ then $ses^{-1} \in \mathcal{U}$ and if, whenever $s \in S$ with $ss^{-1} \in F$, then $s^{-1}s \in F$.*

Proof. From Theorem 2.2.9, the union $\mathcal{U}$ of all the (right) cosets of F is

$$\mathcal{U} = \{s \in S \,:\, ss^{-1} \in F\}.$$

Suppose that F satisfies the conditions in this Theorem. If $s, t \in \mathcal{U}$ then $ss^{-1}, tt^{-1} \in F$ and

$$(st)(st)^{-1} = s(tt^{-1})s^{-1} \in \mathcal{U}$$

which implies that $st \in \mathcal{U}$, and $s^{-1}s \in \mathcal{U}$ which implies that $s^{-1} \in \mathcal{U}$. Hence $\mathcal{U}$ is an inverse subsemigroup.

Conversely, if $\mathcal{U}$ is an inverse subsemigroup and $s \in \mathcal{U}$, then $s^{-1} \in \mathcal{U}$ which implies that $s^{-1}s \in F$, and if $s \in \mathcal{U}$ and $e \in E(F)$ then $e \in \mathcal{U}$ and so $se \in \mathcal{U}$ which implies that $(se)(se)^{-1} = ses^{-1} \in \mathcal{U}$. $\square$

Example 2.2.11. Let $B = \{e, a, a^{-1}, f, y\}$ be a Brandt inverse semigroup, with multiplication table

	e	a	a^{-1}	f	y
e	e	a	y	y	y
a	y	y	e	a	y
a^{-1}	a^{-1}	f	y	y	y
f	y	y	a^{-1}	f	y
y	y	y	y	y	y

and adjoin a zero z to B to obtain S with multiplication table

	e	a	a^{-1}	f	y	z
e	e	a	y	y	y	z
a	y	y	e	a	y	z
a^{-1}	a^{-1}	f	y	y	y	z
f	y	y	a^{-1}	f	y	z
y	y	y	y	y	y	z
z	z	z	z	z	z	z

Then B is a closed inverse subsemigroup of S for which $\mathcal{U} = \{s \in S : ss^{-1} \in B\} = B$. However, if we take $F = \{e\}$ then F is a closed inverse subsemigroup of S for which $\mathcal{V} = \{s \in S : ss^{-1} \in F\} = \{e, a\}$ and here $\mathcal{V}$ is not an inverse subsemigroup of S.

2.3 The Index Formula for Inverse Semigroups

Many significant properties of groups are preserved under passing to subgroups of finite index. There has been a considerable amount of work carried out to establish an analogous concept of index for semigroups in general that preserve properties such as finite generation and finite presentability.

In the theory of semigroups various notions of index have arisen. The first and most widely studied notion of semigroup index has been the so-called Rees index, which was introduced by Jura [12]. The *Rees index* of a subsemigroup is defined simply as the cardinality of its complement. In fact, this index has many nice properties similar to those of group index, and behaves much like the group index. However, the Rees index is not a generalization of the group index. In fact, this notion is very restrictive, and therefore limits the applicability of the results of Rees index. For instance, an infinite group cannot have any proper subgroups of finite Rees index.

Another type of index is the *syntactic index* of Ruškuc and Thomas [25]. It provides a common generalization of group and Rees index. Although, some properties of the group index have their analogues for the syntactic index, this index is too weak. Properties of being finitely generated, finitely presented, and residually finite are not inherited by the syntactic index.

Recently, a new notion of index was introduced, combining both Rees index and group index and at the same time maintaining finiteness conditions, called the *Green index* [9]. The Green index of a subsemigroup H of a semigroup S is defined by counting strong orbits (called H–relative $\mathcal{H}-$ classes) in the complement $S \setminus H$ under the natural actions of H on S by right and left multiplication. Therefore, H has finite Green index in S if there are only finitely many $\mathcal{H}^H$-classes in $S \setminus H$, or, equivalently, if $S \setminus H$ contains only finitely many $\mathcal{R}^H$- and $\mathcal{L}^H$-classes.

In the literature, the notion of index as a measure of the size of H inside S for inverse subsemigroups has barely been investigated. In the current chapter, we begin such an investigation.

Definition The *index* of the closed inverse subsemigroup F in an inverse semigroup S, is the cardinality of the set of right cosets of F, and is written $|\, S \,:\, F \,|$.

Note that the mapping $(Fs)^{\uparrow} \to (s^{-1}F)^{\uparrow}$ is a bijection from the set of right cosets to the set of left cosets.

Definition Let S be an inverse semigroup and let F be a closed inverse subsemigroup. Let us select an element from each right coset of F and write $\mathcal{T}$ for the resulting set. Then we have the union

$$\mathcal{U} \,=\, \bigcup_{t \in \mathcal{T}} (Ft)^{\uparrow},$$

as in Theorem 2.2.9, and every element of $\mathcal{U}$ satisfies $u \geqslant ht$ for some $h \in F,\ t \in T$. The set $\mathcal{T}$ is called a right *transversal* to F in S.

Theorem 2.3.1. *Let S be an inverse semigroup and let H and K be two closed inverse subsemigroups of S with K of finite index in H and H of finite index in S, and with $K \subseteq H$ and K full in S. Then K has finite index in S and*

$$|S : K| = |\, S : H|\, |\, H : K|\,.$$

Proof. Since K is full in S, so is H, and for transversals $\mathcal{T}$, $\mathcal{U}$ we have

$$S = \bigcup_{t \in \mathcal{T}} (Ht)^{\uparrow} \quad \text{and} \quad H = \bigcup_{u \in \mathcal{U}} (Ku)^{\uparrow}.$$

Therefore

$$S \,= \{\, s \in\, S \,:\, s \geqslant ht \quad \textit{for some}\ \ t \in \mathcal{T},\ h \in H \,\},$$

and

$$H \,= \{\, s \in\, S \,:\, s \geqslant ku \quad \textit{for some}\ \ u \in \mathcal{U},\ k \in K \,\}.$$

So if $\ s \geqslant ht\ $ *and* $\ h \geqslant ku$, then $\ s \geqslant kut$. Then $s \in\, (Kut)^{\uparrow}$ and $(Kut)^{\uparrow}$ is a coset of K in S, since K is full in S and therefore $(ut)\,(ut)^{-1} \,\in\, K$.

Hence

$$S = \bigcup_{\substack{u \in \mathcal{U} \\ t \in \mathcal{T}}} (Kut)^{\uparrow}$$

It remains to show that all the cosets $(Kut)^{\uparrow}$ are distinct. Suppose that the opposite is true, that is

$$(Ku't')^{\uparrow} = (Kut)^{\uparrow}.$$

Thus by Proposition 2.2.5(3), $u'\,t'\,t^{-1}\,u^{-1} \in K$. Hence $u'\,t'\,t^{-1}\,u^{-1} \in H$. Since $u \in H$ we have $(u')^{-1}\,u'\,t'\,t^{-1}\,u^{-1}\,u \in H$ and since

$$t'\,t^{-1} \geqslant (u')^{-1}\,u'\,t'\,t^{-1}\,u^{-1}\,u \in H.$$

and H is closed, then $t'\,t^{-1} \in H$. This implies that $(Ht')^{\uparrow} = (Ht)^{\uparrow}$ and therefore $t' = t$.

Now $u'\,t'\,t^{-1}\,u^{-1} \in K$. Since $t' = t$, then $t'\,t^{-1} \in E(S)$. Therefore $u'\,u^{-1} \geqslant u'\,t'\,t^{-1}\,u^{-1}$. But K is closed, so $u'\,u^{-1} \in K$ and $(Ku')^{\uparrow} = (Ku)^{\uparrow}$. Hence, $u' = u$.
Consequently, all the cosets $(Kut)^{\uparrow}$ are distinct. $\square$

Example 2.3.2. Let E be the semilattice of idempotents of S. Recall from Section 1.4 that the property that E is closed is expressed by saying that S is E–unitary. In this case, for any $s \in S$, we have

$$\begin{aligned}(Es)^{\uparrow} &= \{t \in S : t \geqslant es \text{ for some } e \in E\} \\ &= \{t \in S : t \geqslant u \leqslant s \text{ for some } u \in S\} \\ &= \{t \in S : s, t \text{ have a lower bound in } S\}.\end{aligned}$$

We see that $(Es)^{\uparrow}$ is precisely the σ–class of s, where σ is the minimum group congruence on S (see Section 1.3). Obviously $E(S)$ is full in S, every element $t \in S$ lies in a coset of E, and the set of cosets is in one-to-one correspondence with the maximum group image $\widehat{S}$ of S.

Proposition 2.3.3. *Let S be an E–unitary inverse semigroup. Then:*

(1) $E(S)$ has finite index if and only if the maximal group image $\widehat{S}$ is finite, and

$$|S : E| = |\widehat{S}|,$$

(2) if $E(S)$ has finite index in S then, for any closed, full, inverse subsemigroup F of S we have

$$|S : F| = |\widehat{S}| \,/\, |\widehat{F}|$$

Proof. Part (1) follows from our previous discussion. For part (2) we have $E \subseteq F \subseteq S$ and so if the index $|S : E|$ is finite then so are $|S : F|$ and $|F : E|$ with $|S : E| = |S : F||F : E|$. But now $|S : E| = |\widehat{S}|$ and $|F : E| = |\widehat{F}|$. We note that by Lemma 1.6.2, the quotient $|\widehat{S}| \,/\, |\widehat{F}|$ is the index of the subgroup F in S. $\square$

The index formula in Theorem 2.3.1 can still be valid when K is not full in S, as we show in the following example.

Example 2.3.4. The symmetric inverse monoid $\mathscr{I}(\{1,2,3\})$ has 34 elements; where some of them are mentioned in Example 1.9.1.

Take $F = stab(1) = \{\sigma \in \mathscr{I}(X) : 1\sigma = 1\}$. Now

$$F = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & * \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & * & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & * \end{pmatrix}, \right.$$

$$\left. \begin{pmatrix} 1 & 2 & 3 \\ 1 & * & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & * & * \end{pmatrix} \right\}.$$

F is a closed inverse subsemigroup of $\mathscr{I}(X)$ with 7 elements.

First of all, we will find the cosets of F. There are four idempotents in F so if $\sigma\sigma^{-1} \in F$, then the domain of σ is either $\{1,2,3\}$ or $\{1,2\}$ or $\{1,3\}$ or $\{1\}$. It follows that there are 21 possible coset representatives for F in $\mathscr{I}(X)$: six permutations, six with domain $\{1,2\}$, six with domain $\{1,3\}$ and three with domain $\{1\}$. These are listed below:

$$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix},\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix},\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix},\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},$$

$$\begin{pmatrix}1&2&3\\1&2&*\end{pmatrix},\begin{pmatrix}1&2&3\\1&3&*\end{pmatrix},\begin{pmatrix}1&2&3\\3&2&*\end{pmatrix},\begin{pmatrix}1&2&3\\2&1&*\end{pmatrix},\begin{pmatrix}1&2&3\\3&1&*\end{pmatrix},\begin{pmatrix}1&2&3\\2&3&*\end{pmatrix},$$

$$\begin{pmatrix}1&2&3\\1&*&3\end{pmatrix},\begin{pmatrix}1&2&3\\1&*&2\end{pmatrix},\begin{pmatrix}1&2&3\\3&*&1\end{pmatrix},\begin{pmatrix}1&2&3\\2&*&3\end{pmatrix},\begin{pmatrix}1&2&3\\3&*&2\end{pmatrix},\begin{pmatrix}1&2&3\\2&*&1\end{pmatrix},$$

$$\begin{pmatrix}1&2&3\\1&*&*\end{pmatrix},\begin{pmatrix}1&2&3\\2&*&*\end{pmatrix},\begin{pmatrix}1&2&3\\3&*&*\end{pmatrix}.$$

However, these yield only 3 distinct cosets

$$C_1 = \{\sigma \in \mathscr{I}(\{1,2,3\}) :\ 1\sigma = 1\} = F,$$

$$C_2 = \{\sigma \in \mathscr{I}(\{1,2,3\}) :\ 1\sigma = 2\},$$

$$C_3 = \{\sigma \in \mathscr{I}(\{1,2,3\}) :\ 1\sigma = 3\},$$

and so $|\mathscr{I}(\{1,2,3\}) : F\,| = 3$.

Now take

$$K = \left\{\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\right\}.$$

Then K is a closed inverse subsemigroup of F. The domain of each σ in K is $\{1,2,3\}$. and so the only coset representatives for K in F are the permutations

$$\mathrm{id} = \begin{pmatrix}1&2&3\\1&2&3\end{pmatrix} \text{ and } \sigma = \begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}.$$

But these are elements of K and so $(K)^{\uparrow} = (K\sigma)^{\uparrow} = K$ and there is just one coset. Hence

$|F : K| = 1$.

Now, we calculate the cosets of K in $\mathscr{I}(\{1,2,3\})$. Each permutation in $\mathscr{I}(\{1,2,3\})$ is a possible coset representative for K and these produce three distinct cosets by Proposition 2.2.5(3). Hence $|\mathscr{I}(\{1,2,3\}) : K| = 3$ and in this example

$$|\mathscr{I}(\{1,2,3\}) : K| = |\mathscr{I}(\{1,2,3\}) : F|\,|F : K|\ .$$

The index of a closed inverse submonoid L of an inverse monoid M depends on the availability of coset representatives to make cosets, and so on the idempotents of M contained in L. In particular, we can have $L \subset F$ but $|M : L| < |M : F|$ as the following example illustrates.

Example 2.3.5. Consider the free inverse monoid $\mathrm{FIM}(x,y)$ and the closed inverse submonoids $K = \langle x^2 \rangle^{\uparrow}$ and $H = \langle x^2, y^2 \rangle^{\uparrow}$. As in Section 1.11, we represent elements of $\mathrm{FIM}(x,y)$ by Munn trees in the Cayley graph $\Gamma(F(x,y)\{x,y\})$ where $F(x,y)$ is the free group on x, y.

Consider a coset $(K(P,w))^{\uparrow}$. For this to exist, the idempotent $(P,1) = (P,w)(P,w)^{-1}$ must be in K and so as a subtree of Γ, P can only involve vertices in $F(x)$ and edges between them. Since $w \in P$ we must have $w \in F(x)$. It is then easy to see that there are only two cosets, K and $(Kx)^{\uparrow}$ and so $|\,\mathrm{FIM}(x,y) : K| = 2$. Similarly, $|H : K| = 1$.

Now consider a coset $(H(P,w))^{\uparrow}$. Now $(P,1) \in H$ and so P must be contained in subtree of Γ spanned by the vertices of the subgroup $\mathrm{gp}\langle x^2, y^2 \rangle \subset F(x,y)$, and $w \in P$. If $w \in \mathrm{gp}\langle x^2, y^2 \rangle$ then $(P,w) \in H$ and $(H(P,w))^{\uparrow} = H$. Otherwise, $w = ux$ or $w = uy$ with $u \in \mathrm{gp}\langle x^2, y^2 \rangle$ and it follows that there are three cosets of H in $\mathrm{FIM}(x,y)$, namely $H, (Hx)^{\uparrow}$ and $(Hy)^{\uparrow}$. Therefore

$$|\,\mathrm{FIM}(x,y) : H| = 3 > |\,\mathrm{FIM}(x,y) : K| = 2\,.$$

These calculations also follow from results of Margolis and Meakin to be discussed further in Chapter 5, see [18, Lemma 3.2].

We note that the index formula fails to hold. This does not contradict Theorem 2.3.1, since K is not full in $\mathrm{FIM}(x, y)$.

2.4 Actions on Cosets in Inverse Semigroups

Our goal here is to develop enough of the theory of actions on cosets to prove the analogue of M. Hall's Theorem in the next Section. This section is largely based on [26].

Lemma 2.4.1. *For any subset A of S and for any $u \in S$ we have $(Au)^{\uparrow} = ((A)^{\uparrow}u)^{\uparrow}$.*

Proof. Since $A \subseteq (A)^{\uparrow}$ it is clear that $(Au)^{\uparrow} \subseteq ((A)^{\uparrow}u)^{\uparrow}$. On the other hand, if $x \in ((A)^{\uparrow}u)^{\uparrow}$ then for some $a \in A$ we have $y \in S$ with $y \geqslant a$ and $x \geqslant yu$. But then $x \geqslant au$ and so $x \in (Au)^{\uparrow}$. □

Let F be a closed inverse subsemigroup of S, and let $s \in S$ with $ss^{-1} \in F$ with $C = (Fs)^{\uparrow}$. Now suppose that $u \in S$ and that $(Cuu^{-1})^{\uparrow} = C$. Then we define $C \lhd u = (Cu)^{\uparrow}$.

Lemma 2.4.2. *The condition $(Cuu^{-1})^{\uparrow} = C$ for $C \lhd u$ to be defined is equivalent to the condition that $suu^{-1}s^{-1} \in F$.*

Proof. Since for any $c \in C$ we have $c \geqslant cuu^{-1}$ it is clear that $C \subseteq (Cuu^{-1})^{\uparrow}$. Now suppose that $suu^{-1}s^{-1} \in F$ and that $x \in (Cuu^{-1})^{\uparrow} = (Fsuu^{-1})^{\uparrow}$ (by Lemma 2.4.1). Hence there exists $y \in F$ with

$$x \geqslant ysuu^{-1} = yss^{-1}suu^{-1} = ysuu^{-1}s^{-1}s \in Fs.$$

Therefore $x \in (Fs)^{\uparrow} = C$ and so $(Cuu^{-1})^{\uparrow} = C$. On the other hand, if $(Cuu^{-1})^{\uparrow} = C$, then in particular $(Fsuu^{-1})^{\uparrow} \subseteq (Fs)^{\uparrow}$. Since $ss^{-1} \in F$ we have $ss^{-1}suu^{-1} = suu^{-1} \in (Fs)^{\uparrow}$ and so there exists $y \in F$ with $suu^{-1} \geqslant ys$. But then $suu^{-1}s^{-1} \geqslant yss^{-1} \in F$ and since F is closed we deduce that $suu^{-1}s^{-1} \in F$. □

It follows from Lemma 2.4.2 that the condition $suu^{-1}s^{-1} \in F$ does not depend on the choice of coset representative s. This is easy to see directly. If $(Fs)^{\uparrow} = (Ft)^{\uparrow}$ then, by part

(3) of Proposition 2.2.5, we have $st^{-1} \in F$. Then

$$tuu^{-1}t^{-1} \geqslant ts^{-1}suu^{-1}s^{-1}st^{-1} = (st^{-1})^{-1}(suu^{-1}s^{-1})(st^{-1}) \in F$$

and since F is closed, $tuu^{-1}t^{-1} \in F$.

Proposition 2.4.3. *If $u \in S$ and $(Cuu^{-1})^{\uparrow} = C$ then $(Cu)^{\uparrow} = (Fsu)^{\uparrow}$ and the rule $C \lhd u = (Cu)^{\uparrow}$ defines a transitive action of S by partial bijections on the cosets of F.*

Proof. Since $C = (Fs)^{\uparrow}$, Lemma 2.4.1 implies that $(Cu)^{\uparrow} = (Fsu)^{\uparrow}$. To check that $(Fsu)^{\uparrow}$ is a coset of F, we need to verify that $(su)(su)^{-1} \in F$. But $(su)(su)^{-1} = suu^{-1}s^{-1}$ and so this follows from Lemma 2.4.2.

Therefore $(Fsu)^{\uparrow}$ is indeed a coset of F. Moreover, $(Cu)^{\uparrow} \lhd u^{-1}$ is defined and equal to $(Cuu^{-1})^{\uparrow} = C$, so that the action of u is a partial bijection.

It remains to show that for any $s, t \in S$, the action of st is the same as the action of s followed by the action of t whenever these are defined. Now the outcome of the actions are certainly the same: for a coset C, we have $C \lhd (st) = (Cst)^{\uparrow}$ and

$$(C \lhd s) \lhd t = (Cs)^{\uparrow} \lhd t = ((Cs)^{\uparrow}t)^{\uparrow} = (Cst)^{\uparrow},$$

by Lemma 2.4.1.

The conditions for $C \lhd (st)$, $C \lhd s$, $(C \lhd s) \lhd t$ to be defined are, respectively:

$$(Cstt^{-1}s^{-1})^{\uparrow} = C\,, \tag{2.4.1}$$

$$(Css^{-1})^{\uparrow} = C\,, \tag{2.4.2}$$

$$((Cs)^{\uparrow}tt^{-1})^{\uparrow} = (Cstt^{-1})^{\uparrow} = (Cs)^{\uparrow}. \tag{2.4.3}$$

Suppose that (2.4.2) holds. Then

$$(Css^{-1})^{\uparrow} \subseteq (Cstt^{-1}s^{-1})^{\uparrow} = C.$$

But it is clear that $C \subseteq (Css^{-1})^{\uparrow}$, and so $(Css^{-1})^{\uparrow} = C$ and (2.4.2) holds. Now it is again clear that $(Cs)^{\uparrow} \subseteq (Cstt^{-1})^{\uparrow}$, and

$$(Cstt^{-1})^{\uparrow} \subseteq (Cstt^{-1}s^{-1}s)^{\uparrow} = ((Cstt^{-1}s^{-1})^{\uparrow}s)^{\uparrow} = (Cs)^{\uparrow}.$$

Therefore (2.4.3) holds.

Now if both (2.4.2) and (2.4.3) hold we have

$$\begin{aligned}
(Cstt^{-1}s^{-1})^{\uparrow} &= ((Cstt^{-1})^{\uparrow}s^{-1})^{\uparrow} && \text{by Lemma 2.4.1} \\
&= ((Cs)^{\uparrow}s^{-1})^{\uparrow} && \text{by (2.4.3)} \\
&= (Css^{-1})^{\uparrow} && \text{by Lemma 2.4.1} \\
&= C && \text{by (2.4.2).}
\end{aligned}$$

and therefore (2.4.1) holds.

To show that the action is transitive, consider two cosets $(Fa)^{\uparrow}$ and $(Fb)^{\uparrow}$. Then $(Fa)^{\uparrow} \triangleleft a^{-1}b$ is defined since $a(a^{-1}b)(a^{-1}b)^{-1}a^{-1} = aa^{-1}bb^{-1} \in F$, and $(Fa)^{\uparrow} \triangleleft a^{-1}b = (Faa^{-1}b)^{\uparrow} = (Fb)^{\uparrow}$, since again $aa^{-1}bb^{-1} \in F$. $\square$

2.5 An Inverse Semigroup Version of M. Hall's Theorem

Finite generation arise in many contexts of semigroup theory in general. In order to describe some of its connections with inverse semigroups, it will be convenient to recall some of the theory of finite generated groups. For further background information, see [3, 24].

Theorem 2.5.1. *(M. Hall's Theorem for Groups) In a finitely generated group G, there exist only finitely many distinct subgroups of G having a fixed finite index d. In fact, if G has a generating set of size m, then the number of possible subgroups of index d is at most $(d!)^m$.*

Proof. Suppose that the group G is generated by m generators. Let H be a subgroup of index d. Take a transversal $\{t_1, t_2, \ldots, t_d\}$, so any coset $Hg = Ht_j$ for a unique j. Choose $t_1 = e_G$, so that $Ht_1 = H$. Then if $Ht_i \lhd g = Ht_j$, we get a permutation of $\{1, 2, \ldots, d\} = X_d$ by defining $i \lhd g = j$.
This gives a homomorphism $\phi_H : G \longrightarrow \Sigma_d$ into the symmetric group of all permutations of X_d.

Suppose $g \in G$ satisfies $1 \lhd g = 1$, so $g \in \mathrm{stab}(1) = \{g \in G : 1 \lhd g = 1\}$. This is equivalent to $Ht_1 \lhd g = Ht_1$. This implies that

$$g \in \mathrm{stab}(1) \Longleftrightarrow Hg = H \Longleftrightarrow g \in H\,.$$

Hence $\mathrm{stab}(1) = H$, and it follows that H is determined by ϕ_H. Therefore, the number of subgroups of index d is at most the number of homomorphisms

$$\phi_H : G \longrightarrow \Sigma_d\,.$$

But this is finite, since ϕ is determined by a choice of image for each generator of G, and so we have m choices to make from $d!$ elements of Σ_d. This gives an upper bound $(d!)^m$, as stated above. $\square$

Theorem 2.5.2. *(M. Hall's Theorem for Inverse Semigroups) In a finitely generated inverse semigroup S there are at most finitely many distinct closed inverse subsemigroups of a fixed finite index d.*

Proof. Suppose that the inverse semigroup S is finitely generated and that the closed inverse subsemigroup F of S has exactly d cosets. We aim to construct an inverse semigroup homomorphism

$$\phi_F : S \longrightarrow \mathscr{I}(X),$$

where $\mathscr{I}(X)$ is the symmetric inverse monoid on $X = \{1, ..., d\}$.

Write the distinct cosets of F as $(Fc_1)^{\uparrow}, (Fc_2)^{\uparrow}, \ldots, (Fc_d)^{\uparrow}$, with $c_1, c_2, ..., c_d \in S$, and with $(Fc_1)^{\uparrow} = F$. Now take $u \in S$. If $c_j\, u\, u^{-1}\, {c_j}^{-1} \in F$, where $j \in \{1, ..., d\}$, then we can define the action of the element $u \in S$ on the coset $(Fc_j)^{\uparrow}$ of F as follows:

$$(Fc_j)^{\uparrow} \lhd u = (Fc_j u)^{\uparrow}.$$

By Proposition 2.4.3, $(Fc_j u)^{\uparrow}$ is indeed a coset of F, and so $(Fc_j u)^{\uparrow} = (Fc_k)^{\uparrow}$, where $k \in \{1, ..., d\}$. Then we can write $(Fc_j)^{\uparrow} \lhd u = (Fc_k)^{\uparrow}$. By Proposition 2.4.3, this action of u defines an action of S by partial bijections on the cosets of F, and therefore we do get the homomorphism

$$\phi_F : S \longrightarrow \mathscr{I}(X).$$

We now need to show that different choices of F give us different homomorphisms ϕ_F. In other words, to establish the claim that $\phi_F = \phi_K$ then $F = K$, we have to reconstruct F from ϕ_F. In order to do this, if we claim that if $x \in S$ and $x\phi_F$ acts on 1, so that $1 \lhd x\phi_F = 1$, then $x \in F$, and conversely.

By Lemma 2.4.2, for any coset C and $u \in S$, $C \lhd u = (Cu)^{\uparrow}$ if and only if $(Cuu^{-1})^{\uparrow} = C$. If $C = (Fs)^{\uparrow}$, then $C \lhd u = (Fsu)^{\uparrow}$ if and only if $suu^{-1}s^{-1} \in F$. We specialize to the case $C = F$. Firstly, we show that If $x \in F$, we prove that $F \lhd x = F$. Suppose $x \in F$, then $x\,x^{-1} \in F$, and so $Fxx^{-1} \subseteq F$ and therefore $(Fxx^{-1})^{\uparrow} \subseteq (F)^{\uparrow} = F$. But also, if $y \in F$, then $y \geqslant y\,x\,x^{-1}$, and so $y \in (Fxx^{-1})^{\uparrow}$. Therefore $(Fxx^{-1})^{\uparrow} = F$. Thus $F \lhd x$ is defined and equals $(Fx)^{\uparrow}$. But $Fx \subseteq F$, so $(Fx)^{\uparrow} \subseteq (F)^{\uparrow} = F$. But also, if $y \in F$, then $y \geqslant y\,x^{-1}x \in Fx$ and $y \in (Fx)^{\uparrow}$. Hence $(Fx)^{\uparrow} = F$, and so that $F \lhd x = F$. Now if $F \lhd x = F$, we show that $x \in F$.

Suppose that $F \lhd x$ is defined, so $(Fxx^{-1})^{\uparrow} = F$. If $e \in E(F)$, then $xx^{-1} \geqslant exx^{-1} \in Fxx^{-1}$ and so $xx^{-1} \in F$. Then $x = x\,x^{-1}x \in F \lhd x = F$. Therefore $x \in F$. Hence $\mathrm{stab}(F) = F$ and in the induced action of S on X we have $\mathrm{stab}(1) = F$, and F is determined by ϕ_F.

This implies that if we have two homomorphisms ϕ_{F_1}, ϕ_{F_2} (derived from the inverse subsemigroups F_1, F_2 respectively), and we know that $\phi_{F_1} = \phi_{F_2}$, then $F_1 = F_2$. Therefore, the number of closed inverse subsemigroups of index d is at most the number of homomorphisms $\phi : S \longrightarrow \mathscr{I}(X)$.

Since S is finitely generated; say it has n generators, and $\mathscr{I}(X)$ is finite, therefore there are only finitely many homomorphisms $\phi : S \longrightarrow \mathscr{I}(X)$, and so we can conclude that the set Φ of all such homomorphisms is finite. More precisely, ϕ is determined by a choice of image for each generator of S, so we have n choices to make from $|\mathscr{I}(X)|$ elements of $\mathscr{I}(X)$. Then we obtain

$$|\,\Phi\,| \leqslant |\,\mathscr{I}(X)\,|^n$$

However, by the result of Theorem 1.9.6, we know that the size of $\mathscr{I}(X)$ is

$$|\,\mathscr{I}(X)\,| = \sum_{r=0}^{d} \binom{d}{r}^2 r!\,.$$

Hence, the number of inverse subsemigroups of S of index d is at most $|\,\Phi\,|$, where

$$|\,\Phi\,| \leqslant \left(\sum_{r=0}^{d} \binom{d}{r}^2 r! \right)^n .$$

This completes the proof that there exist only finitely many distinct closed inverse subsemigroups of S having a fixed finite index d. $\square$

Chapter 3

Cosets in the Bicyclic Monoid

In this chapter, our attention is primarily on the theory of the bicyclic monoid in the first half and then later on Bruck-Reilly extensions. We establish the foundations of the theory of closed inverse subsemigroups of the bicyclic monoid and then determine the index of these closed inverse subsemigroups. We will also see that understanding the structure of the bicyclic monoid B and its closed inverse subsemigroups is the key to understanding other more general structures, namely the Bruck-Reilly extensions. Section 3.5 is devoted to the construction of Bruck-Reilly semigroups and then describes the index of the copy of the bicyclic monoid B in Bruck-Reilly semigroups in various settings.

3.1 Definitions and Preliminaries

The *bicyclic monoid* B is defined by the monoid presentation $\langle b, c : bc = 1\rangle$, is one of the most fundamental semigroups. It is an example of an infinite but finitely generated inverse monoid with an infinite descending chain of idempotents $1 > cb > c^2b^2 > c^3b^3 > \dots$. Clearly in B, we have $c = b^{-1}$ and so we may also define B by the presentation

$$B = \langle b \mid bb^{-1} = 1\rangle . \tag{3.1.1}$$

The natural description of its set of elements as $\{ c^i b^j \,:\, i, j \geqslant 0 \}$ and so we can describe the elements of B as pairs of non-negative integers

$$B = \{ (i, j) \,:\, i, j \geqslant 0 \}$$

and the multiplication rule for these forms is according to the following:

$$(m, n)\,(r, s) = \begin{cases} (m, n - r + s) & \text{if} \quad n \geqslant r \ ; \\ \\ (m - n + r, s) & \text{if} \quad n \leqslant r \ . \end{cases} \tag{3.1.2}$$

3.2 Basic Properties

The main structural features of the bicyclic monoid are provided in the following result from [14]. We write

$$a \ominus b = \begin{cases} a - b & \text{if} \quad a \geqslant b \ ; \\ 0 & \text{otherwise} \, . \end{cases}$$

On the set $\mathbb{N} \times \mathbb{N}$, define a binary operation by

$$(m, n)\,(r, s) = (m + (r \ominus n), s + (n \ominus r)). \tag{3.2.1}$$

Theorem 3.2.1. *([14] Theorem 3.4.3) The bicyclic monoid is a combinatorial, bisimple, $E-$unitary inverse monoid.*

Proof. In order to find the idempotents in B, we want to find the elements (m, n) such that $(m, n)\,(m, n) = (m, n)$. Now consider the case when $m \geqslant n$

$$\begin{aligned} (m, n)(m, n) &= (m + (m \ominus n), n + (n \ominus m)) \\ &= (m + m - n, n + 0) \\ &= (2m - n, n) \end{aligned}$$

But $(2m-n,n)=(m,n)$, so that $2m-n=m$ and then $m=n$.

In the case when $n \geqslant m$, we get

$$\begin{aligned}(m,n)\,(m,n) &= (m+0, n+n-m)\\ &= (m, 2n-m)\end{aligned}$$

Thus $2n-m=n$ and then $m=n$. Hence, in both cases the idempotents are of the form (m,m). Suppose that (m,m) and (n,n) are idempotents. Consider these two cases:

- If $m \geqslant n$, then $(m,m)\,(n,n)=(m+0, m-n+n)=(m,m)$ and $(n,n)\,(m,m)=(n+m-n, m+0)=(m,m)$.
- If $n \geqslant m$, then $(m,m)\,(n,n)=(m+n-m, n+0)=(n,n)$ and $(n,n)(m,m)=(n+m-n, m+0)=(n,n)$.

Hence $(m,m)\,(n,n)=(n,n)(m,m)$. It follows that the idempotents in B commute and thus the idempotents form a commutative subsemigroup. To show that B is regular, observe that $(m,n)\,(n,s)=(m+(n\ominus n), s+(n\ominus n))=(m+0,s+0)=(m,s)$. Thus $(m,n)(n,s)=(m,s)$. This implies that $(m,n)(n,m)(m,n)=(m,m)(m,n)=(m,n)$. Thus the bicyclic monoid is regular with commuting idempotents, and so B is an inverse monoid. Notice that

$$(n,m)(m,n)(n,m)=(n,n)(n,m)=(n,m)\,.$$

Therefore the inverse of (m,n) is (n,m). Moreover,

$$(m,n)^{-1}(m,n)=(n,n) \quad \text{and} \quad (m,n)(m,n)^{-1}=(m,m)\,.$$

Therefore, the bicyclic monoid is combinatorial.

Take two idempotents (m,m) and (n,n) then the element (m,n) satisfies $(m,m)\,\mathcal{R}\,(m,n)$ and $(m,n)\,\mathcal{L}\,(n,n)$. Hence $(m,m)\,\mathcal{D}\,(n,n)$ and so the bicyclic monoid is bisimple.

Next, we characterize the natural partial order. Assume that $(m,n) \leqslant (p,q)$. Then $(m,n) = (p,q)\,(n,n)$. Now using the definition of the binary operation in (3.2.1), we get

$$(p,q)(n,n) = (p + (n \ominus q), n + (q \ominus n) = (m,n)\,.$$

Thus

$$m = p + (n \ominus q) \quad \text{and} \quad n = n + (q \ominus n)\,.$$

But $n = n + (q \ominus n)$ implies that $q \ominus n = 0$ and then $q \leqslant n$. Put $a = n \ominus q$. Then $m = a + p$ and $n = a + q$. Conversely, assume that $m = a + p$ and $n = a + q$, for some $a \in \mathbb{N}$. Therefore

$$\begin{aligned}
(p,q)\,(n,n) &= (p + (n \ominus q), n + (q \ominus n) \\
&= (p + n - q, n + o) \quad \text{(since } n \geqslant q) \\
&= (p + a, n) \\
&= (p + a, n) \quad \text{(since } m = a + p) \\
&= (m,n)
\end{aligned}$$

Therefore $(m,n) \leqslant (p,q)$. We have therefore proved that the natural partial order on the bicyclic monoid is given by:

$$(m,n) \leqslant (p,q) \Longleftrightarrow m = a + p \ \text{ and } \ n = a + q\,,$$

for some $a \in \mathbb{N}$.

Suppose that $(m,m) \leqslant (p,q)$. Then $m = p+a$ and $m = q+a$. So, we have $p+a = q+a$ and then $p = q$. Hence (p,q) is an idempotent. It implies that any element above an idempotent in the natural partial order is an idempotent as well. Therefore, the bicyclic monoid B is $E-$unitary. □

It is often convenient to view B as an infinite square grid. The following Figures show the structure of the bicyclic monoid B.

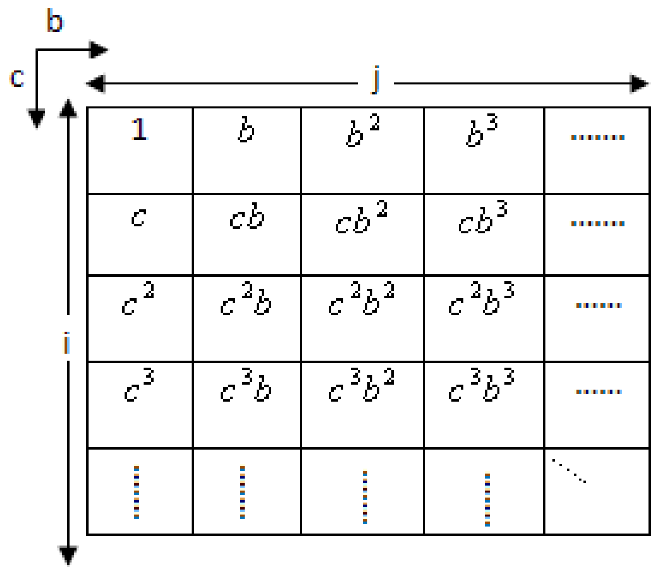

Figure 1: *The Bicyclic monoid* B

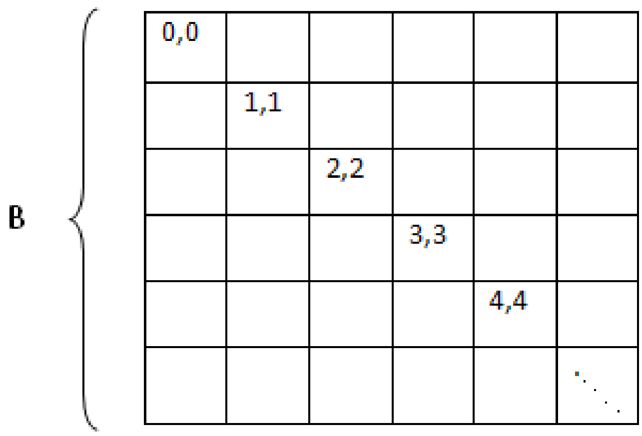

Figure 2: *The semilattic of idempotents* $E(B)$

3.3 Closed Inverse Subsemigroups of B

The main focus of this section is on describing all the closed inverse subsemigrops of the bicyclic monoid B, and classify them according to their index, whether it's finite or infinite. Lemma 3.3.5 and Theorem 3.3.6 give us a full classification of the closed inverse subsemigroups of B. In Chapter 4, we shall generalize the ideas of the proof to consider closed inverse subsemigroups of the polycyclic monoids (following Lawson [16]) and of graph inverse semigroups.

We now introduce some notations for certain subsets of B, taken from [7], in which Descalço and Ruškuc undertook a detailed study of the subsemigroups of B.

$D = \{\,(p,p)\;:\;p \in \mathbb{N}\,\}$ – called the *diagonal* of B,

$M_d \;=\; \{\,(p,q)\;:\;q \equiv p \pmod{d}\}\,,$

$\Sigma_{p,d,P} = \{\,(p+r+ud, p+r+vd)\;:\;r \in P,\; u,v \geqslant 0\,\}$ – the *squares* of length d where $p \in \mathbb{N},\; p \geqslant 0,\; d > 0,\; [d] = \{\,0,\ldots,d-1\,\}$ and $P \subseteq [d]$.

Analogously, it seems to be more useful in order to achieve our results to represent B as an infinite square grid; as we have done in the last section in Figure 1, and then define its subsets, as shown in the Figures below. Some examples illustrating these subsets of B are provided in the following Figures 3 and 4 below.

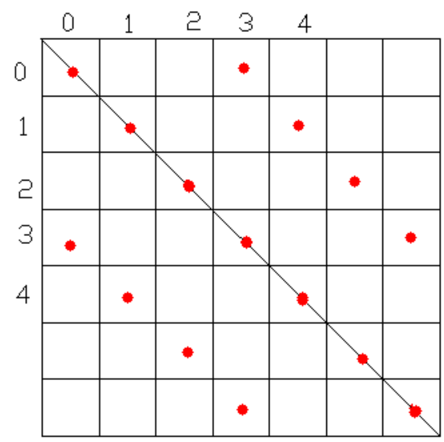

Figure 3: *The $\lambda-$ multiples of 3, M_3*

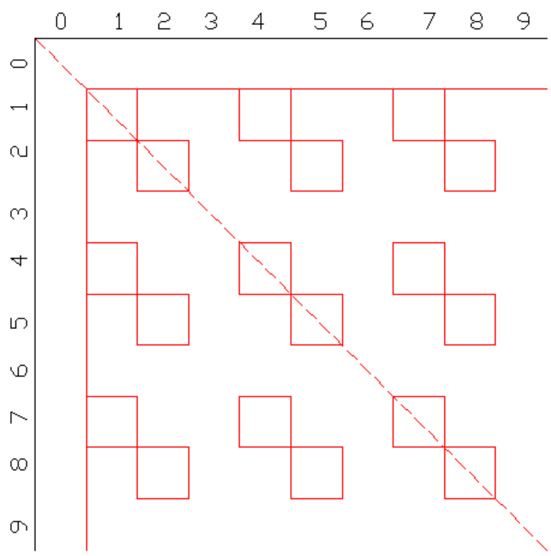

Figure 4: *The collection of squares* $\sum_{1,3,\{0,1\}}$ *of length* 3

Another way of describing these squares is as follows.

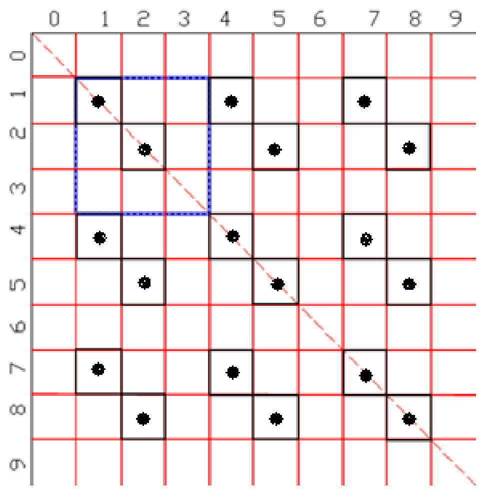

Figure 5: *The structure of* $\sum_{1,3,\{0,1\}}$

The general picture for these families of squares $\sum_{p,d,P}$ of length d for any positive integer $d > 0$ can be viewed in a fairly simple sketch provided in Figure 6.

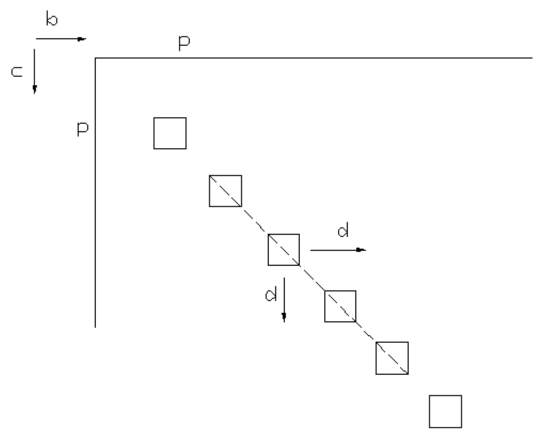

Figure 6: *The squares* $\sum_{p,d,P}$

We will provide some significant results from the work of Descalço and Ruškuc [7] without proof, which are of great relevance to our purpose in this chapter.

Lemma 3.3.1. *For any* $p \in \mathbb{N}_0,\ d \in \mathbb{N},$ *and* $P \subseteq \{0, \dots, d-1\},$ *the square* $\sum_{p,d,P}$ *is a subsemigroup.*

Theorem 3.3.2. *A subsemigroup* $\mathcal{S}$ *of* B *is regular (and hence inverse) if and only if it has the form* $F_D \bigcup \sum_{p,d,P}$ *where* F_D *is a finite subset of the diagonal and either of* F_D *or* P *may be empty.*

Remark 3.3.3. It has been shown in Ruškuc's work [7] that K; the minimal two-sided ideal of the inverse subsemigroup S of B, is isomorphic to M_d. In the context of this Theorem 3.3.2, we have that $C = F_D$ and $K = \sum_{p,d,P}$. Thus M_d is isomorphic to $\sum_{p,d,P}$, and therefore $\sum_{p,d,r}$ is isomorphic to B.

Another important result we will provide here is

Corollary 3.3.4. *A subsemigroup* $\mathcal{S}$ *of* B *is regular (and hence inverse) if and only if it is obtained by adjoining successively finitely many identities to a finite union of copies of* B*, in other words, to* M_d *.*

Now, we describe in the next Lemma 3.3.5, all the closed inverse subsemigroups of the semillatice of idempotents of B.

Lemma 3.3.5. *The only closed inverse subsemigroups of* $E(B)$ *are:*

$E(B)$ *itself and* $E(B)_N = \{(p,p) : p \leqslant N\}$ *for any* $N \in \mathbb{N}$.

Proof. We note that if $(p,p) \in E(B)$ then $(q,q) \geqslant (p,p) \iff q \leqslant p$. So, if a closed inverse subsemigroup of $E(B)$ contains (p,p), then it contains (q,q) for all $q \leqslant p$. Moreover, since

$$(p,p)(q,q) = (\max\{p,q\}, \max\{p,q\})$$

it is clear that $E(B)_N$ is a closed inverse subsemigroup of $E(B)$.

Now if F is a closed inverse subsemigroup of $E(B)$, but is not equal to $E(B)$ then it contains an element (N,N) with N maximal: then by closure, $F = E(B)_N$. ☐

The closed inverse subsemigroups of $E(B)$ are illustrated in Figure 7.

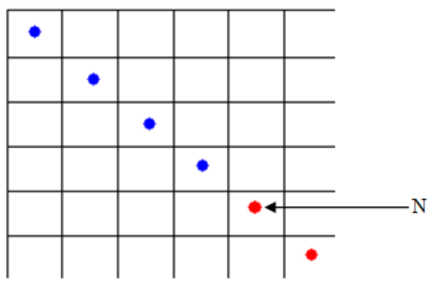

Figure 7: *The closed inverse subsemigroup* $E(B_N)$

We observe that the natural partial order on $E(B_N)$ is a finite chain of idempotents: $e_1 > e_2 > ... > e_n$ and $e_i\, e_j = e_{max(i,j)}$, as shown in Figure 8.

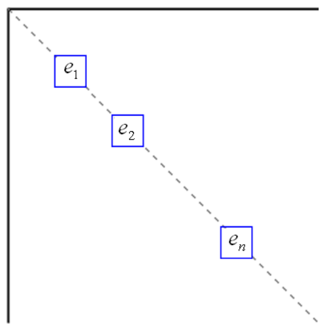

Figure 8: *Finite inverse subsemigroup* $E(B_N)$

Theorem 3.3.6. *An inverse subsemigroup* L *of* B*, not contained in* $E(B)$*, is closed if and only if it has the form* $M_d = \{ (p,q) : q \equiv p \pmod{d} \}$ *for some* $d > 0$*.*

Proof. Let L be a closed inverse subsemigroup of the bicyclic monoid

$$B = \{ a^{-p}\, a^{q} : p, q \geqslant 0 \}, \quad aa^{-1} = 1 \ .$$

If $L \subseteq E(B)$, then by Lemma 3.3.5, L can be a finite chain

$$L = \{ a^{-k}\, a^{k} : 0 \leqslant k \leqslant m \},$$

for some m, or can be $L = E(B)$, an infinite chain of idempotents.

If $L \nsubseteq E(B)$, we can find $a^{-p}a^{q} \in L$ with $q - p > 0$ minimal.
Hence $a^{-p}a^{q} = a^{-p}a^{p+(q-p)} \in L$ and, since L is closed, we have $a^{q-p} \in L$. (Recall that in B the natural partial order in B is as follows:

$$\begin{aligned} a^{-k}a^{l} &\geqslant a^{-n}\, a^{n}a^{-k}a^{l} \\ &= a^{-n}a^{n-k+l} \quad (n \geqslant k) \\ &= a^{-k}a^{l} \qquad\quad (k \geqslant n) \end{aligned}$$

In the first case, write $n = k + r$, then

$$a^{-k}a^{l} \geqslant a^{-(k+r)}a^{r+l},$$

so we reduce each power by r to move up the natural partial order.)

Set $q - p = d$ so that $a^{d} \in L$. Since L is a subsemigroup, this implies that $a^{nd} \in L$, for any $n \geqslant 1$, so powers of a occurring in L can be arbitrarily large. Hence the idempotent $a^{-nd}a^{nd} \in L$ associated to $a^{nd} \in L$ involves an arbitrarily large exponent. Since L is upwardly closed, then L includes every idempotent above $a^{-nd}a^{nd}$ and so L is full.

Now we define

$$M_d = \{a^{-p}a^{q} : \ q - p \ \text{ divisible by } d\}.$$

Then for any $r \leqslant \min\{p,q\}$, we have

$$a^{-(p-r)}a^{q-r} \geqslant a^{-p}\,a^{q} \in M_d,$$

and $(q-r-p+r) = q-p$ is divisible by d. Therefore M_d is a closed inverse subsemigroup of B.

Take an element $a^{-p}a^{q} \in M_d$. Since $q-p$ is divisible by d, then we can write $q-p = dk$, for some integer k. Hence

$$\begin{aligned} a^{-p}a^{q} &= a^{-p}a^{p+dk} \\ &= a^{-p}a^{p}(a^{d})^{k}\,. \end{aligned}$$

Since L is full, then $a^{-p}a^{p} \in L$, and since by assumption $a^{d} \in L$, then $(a^{d})^{k} \in L$. Hence, $a^{-p}a^{q} \in L$ and therefore $M_d \subseteq L$.

Now we need to prove that $L \subseteq M_d$. Suppose that $a^{-s}a^{t} \in L$, where $t \geqslant s$. Then since L is closed, $a^{t-s} \in L$. Assume that $t-s$ is not divisible by d. Write $t-s = dk+r$, where $0 \leqslant r < d$. Then

$$\begin{aligned} a^{r} &= a^{t-s-dk} \\ &= a^{t-s}(a^{d})^{-k}\,. \end{aligned}$$

Since $(a^{d})^{-k} \in L$ and $a^{t-s} \in L$ and L is an inverse subsemigroup, then $a^{r} = a^{t-s}(a^{d})^{-k} \in L$. Minimality of d then implies that $r = 0$, and so that d divides $t-s$, and therefore $a^{-s}a^{t} \in M_d$. Hence $L \subseteq M_d$, and so $L = M_d$. Consequently, any closed inverse subsemigroup of B is of this form

$$M_d = \{\,(p,q) :\, p \equiv q \pmod{d}\}\,.$$

$\square$

We show some concrete examples of M_d:

Example 3.3.7. If $d = 1$, then: $M_1 = \{(p, q) : q \equiv p \pmod 1\} = B$.

If $d = 2$, then $M_2 = \{(m, n) : m \equiv n \pmod 2)\}$. It seems to be sufficient to illustrate the situation using the square grid below.

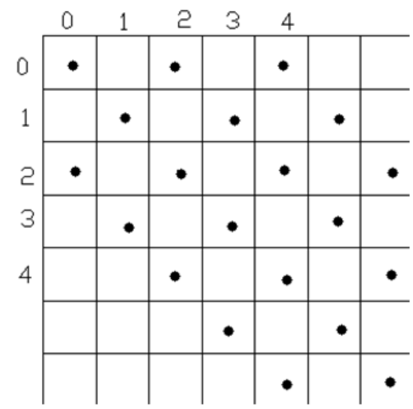

Figure 9: *The inverse subsemigroup* M_2

If $d = 3$, we also can represent M_3 by the form $M_3 = \{(m, n) : m \equiv n \pmod 3)\}$. Analogously, this process can be done for any value of $d > 0$.

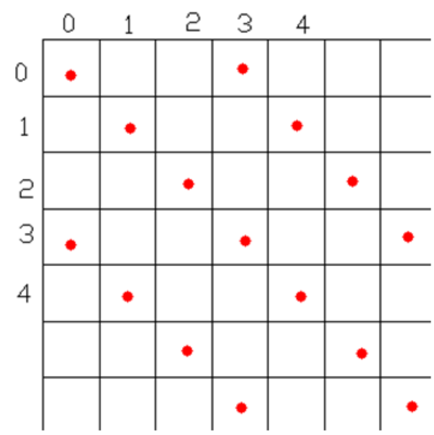

Figure 10: *The inverse subsemigroup* M_3

The inverse subsemigroups M_d are discussed in detail in Chapter 5 of [11] (where they are denoted B_d). It is shown in [11, Proposition 5.7.5] that the M_d are, up to isomorphism, the only fundamental inverse ω–semigroups.

It is in the light of the previous results, Lemma 3.3.5 and Theorem 3.3.6, that we can state and prove the next results regarding the index of closed inverse subsemigroups in the bicyclic monoid B.

Theorem 3.3.8. *The semilattice of idempotents $E(B)$ has infinite index in B, and so does each closed inverse subsemigroup $E(B)_N$.*

Proof. It is clear from the presentation (3.1.1) that $\widehat{B} \cong \mathbb{Z}$, and so by Proposition 2.3.3, $E(B)$ has infinite index in $\mathbb{Z}$.

Now fix $N \geqslant 0$, and consider two cosets $(E(B)_N\,(m,p))^{\uparrow}$ and $(E(B)_N\,(n,q))^{\uparrow}$, where $0 \leqslant m \leqslant N,\ 0 \leqslant n \leqslant N$, and $p \geqslant q \geqslant 0$. These two cosets are equal if and only if

$$(m,p)(q,n) = (m, p-q+n) \in E(B)_N,$$

that is, if and only if $m - n = p - q$. If $m = n$ then $p = q$, and so if $p \neq q$ the cosets $(E(B)_N\,(m,p))^{\uparrow}$ and $(E(B)_N\,(n,q))^{\uparrow}$ are distinct, and hence $E(B)_N$ has infinite index in B. $\square$

Theorem 3.3.9. *A typical coset of the closed inverse subsemigroup M_d in B has the form:*

$$(M_d\,(0,k))^{\uparrow} = \{(p,q) :\ -p+q \equiv k \pmod{d}\}$$

for some $k \in \{0,1,\ldots,d-1\}$. Moreover, the closed inverse subsemigroup M_d has finite index d in B.

Proof. It is clear that M_d is full in B, and so we observe that each element of B determines a coset. Now an element $(p,q) \in B$ is in the coset $(M_d\,(0,k))^{\uparrow}$ if and only if $(p,q)(k,0) \in M_d$. But

$$(p,q)(k,0) = \begin{cases} (p, q-k) & \text{if} \quad q \geqslant k\,, \\ (p-q+k, 0) & \text{if} \quad q < k\,. \end{cases}$$

Now $(p, q-k) \in M_d$ if and only if $-p+q-k$ is divisible by d, that is, if and only if $-p+q \equiv k \pmod{d}$. Similarly, $(p-q+k, 0) \in M_d$ if and only if d divides $-p+q-k$, that is, if and only if $-p+q \equiv k \pmod{d}$. This confirms the description of the coset

$$(M_d\,(0,k))^{\uparrow} = \{(p,q) :\ -p+q \equiv k \pmod{d}\}.$$

Since M_d is full and closed, the union of its cosets is equal to B, and since an arbitrary element (p,q) belongs to exactly one coset, we see that there are d distinct cosets $(M_d\,(0,k))^{\uparrow}$ for $k \in \{0,1,\ldots,d-1\}$ and hence that M_d has index d in B. $\square$

Next, we provide examples to show the nature of the cosets of some closed inverse subsemigroups M_d describing exactly how many there are and their forms.

Example 3.3.10. There are two cosets of the inverse subsemigroup M_2 as follows:

$$(M_2\,(0,k))^{\uparrow} = \{(p,q) :\ -p+q \equiv k \pmod{2}\}$$

where $k \in \{0,1\}$. Therefore, we get two cosets:

$$M_2 = (M_2)^{\uparrow} = \{(p,q) :\ -p+q \equiv 0 \pmod{2}\},$$

$$M_2(0,1)^{\uparrow} = \{(p,q) :\ -p+q \equiv 1 \pmod{2}\}.$$

Next example shows the cosets of the inverse subsemigroup M_4 of B.

Example 3.3.11. There are exactly four cosets of the inverse subsemigroup M_4 in B. They are shown in the following Figure:

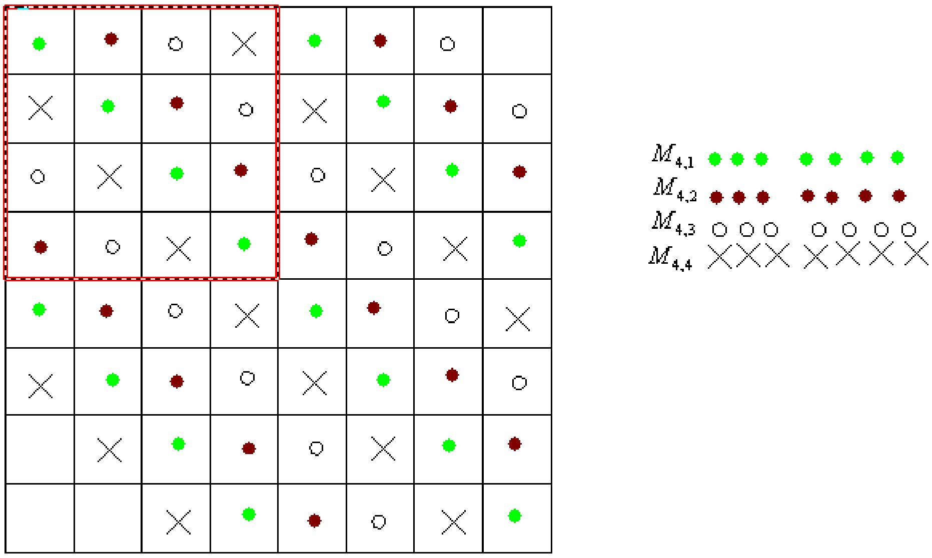

Figure 11: *The cosets of the inverse subsemigroup* M_4

3.4 Generating Sets for Finite Index Inverse Subsemigroups of The Bicyclic Monoid

We have established a full description for finite index inverse subsemigroups of the bicyclic monoid and hence justified our discussion of the exact number of the index. Our next step now is to prove that these inverse subsemigroups are finitely generated and to further understand their structure.

Lemma 3.4.1. *The closed inverse subsemigroup M_d of B is generated by the elements*

$$(0, d),\ (1, 1+d),\ \ldots\ (d-1, 2d-1)\,.$$

Proof. We construct for each element $(a+rd, a+sd) \in M_d$ an explicit representation of it, in terms of these generators. We can assume that $a \in \{0, 1, \ldots, d-1\}$.

Now $(a, a+d)^2 = (a, a+d)\,(a, a+d) = (a, a+2d)$ and then inductively, we see that $(a, a+d)^s = (a, a+sd)$. Similarly, $(a+d, a)^2 = (a+d, a)\,(a+d, a) = (a+2d, a)$ and

$(a+d,a)^r = (a+rd,a)$. Hence

$$(a+d,a)^r\,(a,a+d)^s = (a+rd,a)\,(a,a+sd) = (a+rd,a+sd)\,.$$

$\square$

Theorem 3.4.2. *The closed inverse subsemigroup M_d is a disjoint union of d isomorphic copies of B, where each copy of B is an inverse subsemigroup of M_d (and of B) generated by $(a,a+d)$, for some a with $0 \leqslant a \leqslant d-1$.*

Proof. We write $\langle(a,a+d)\rangle$ for the inverse subsemigroup of M_d generated by $(a,a+d)$. We study, first how the multiplication works in $\langle(a,a+d)\rangle$. Put $\alpha = (a,a+d)$, and then

$$\begin{aligned}
(a,a+d)\,(a+d,a) &= (a,a)\\
(a,a)\,(a,a+d) &= (a,a+d)\\
(a+d,a)\,(a,a) &= (a+d,a)\\
(a,a+d)\,(a,a) &= (a,a+d)\\
(a,a)\,(a+d,a) &= (a+d,a)\,.
\end{aligned}$$

So (a,a) behaves like the element $(0,0)$ in $B;$ it is the multiplicative identity for $\langle(a,a+d)\rangle; 1_{\langle(a,a+d)\rangle}$. Hence $\alpha\alpha^{-1} = 1_{\langle(a,a+d)\rangle}$.

We know that a typical element of $\langle(a,a+d)\rangle$ is a string of α's and α^{-1}'s which collapses to α^{-r} and α^{s} for some r,s. Thus $\alpha^s = (a,a+sd)$, also $\alpha^r = (a,a+rd)$, and so $\alpha^{-r} = (a+rd,a)$. Hence

$$\begin{aligned}
\alpha^{-r}\,\alpha^{s} &= (a+rd,a)\,(a,a+sd)\\
&= (a+rd,a+sd)\,.
\end{aligned}$$

So we now know that

$$\langle(a,a+d)\rangle = \{(a+rd,a+sd) :\ r,s \geqslant 0\}\,.$$

We now show that the mapping θ defined on a typical element in M_d by

$$\theta : (a + rd, a + sd) \longmapsto (r, s)$$

gives an isomorphism

$$\langle (a, a + d) \rangle \to B.$$

Clearly, θ is a bijection, since each element in the bicyclic monoid B corresponds to a unique element of $\langle (a, a + d) \rangle$.

Take $x_1 = (a + r_1 d, a + s_1 d)$, $x_2 = (a + r_2 d, a + s_2 d)$.

Now

$$\begin{aligned} x_1\, x_2 &= (a + r_1 d, a + s_1 d)\,(a + r_2 d, a + s_2 d) \qquad \text{(assuming } a + s_1 d \leqslant a + r_2 d\text{)} \\ &= (a + r_1 d - a - s_1 d + a + r_2\, d, a + s_2 d) \quad \text{(by the multiplication rule on } B\text{)} \\ &= (a + (r_1 - s_1 + r_2)\, d, a + s_2 d)\ . \end{aligned}$$

By applying the map θ, we get

$$(x_1\, x_2)\, \theta = (r_1 - s_1 + r_2, s_2)\ .$$

Also

$$\begin{aligned} x_1\, \theta\ .\ x_2\, \theta &= (a + r_1 d, a + s_1 d)\ \theta\ .\ (a + r_2 d, a + s_2 d)\ \theta \\ &= (r_1, s_1)\,(r_2, s_2) \qquad \text{(by the definition of the map } \theta\text{)} \\ &\qquad \text{(assuming } a + s_1 d \leqslant a + r_2 d \Leftrightarrow s_1 \leqslant r_2\text{)} \\ &= (r_1 - s_1 + r_2, s_2)\ . \quad \text{(by the multiplication rule on } B\text{)} \end{aligned}$$

Therefore

$$(x_1\, x_2)\, \theta = x_1\, \theta\ .\ x_2\, \theta,$$

and so θ is a homomorphism. Consequently, the mapping θ is an isomorphism and so we proved the claim that $\langle(a, a+d)\rangle$ is isomorphic to B.

This implies that the closed inverse subsemigroup M_d is a disjoint union of d copies of B.

$$M_d \cong \underbrace{B \cup B \cup \ldots \cup B}_{d \ \ copies} \ .$$

□

This leads us to the following result:

Theorem 3.4.3. *M_d has d distinct $\mathcal{D}$-classes.*

Proof. We know from Theorem 3.4.2, that M_d is a disjoint union of d copies of B, so

$$M_d \cong \underbrace{B \cup B \cup \ldots \cup B}_{d \ \ copies} \ .$$

Also, B has the property that it is bisimple, see Theorem 3.2.1, if and only if Green's $\mathcal{D}$-relation has one class. Therefore, each copy of B is a $\mathcal{D}$- class in M_d. It follows that M_d has d distinct $\mathcal{D}$-classes. □

The next result follows immediately from the above.

Corollary 3.4.4. *If $d \neq d'$, then M_d is not isomorphic to $M_{d'}$.*

The next example is to illustrate the idea in Theorem 3.4.2.

Example 3.4.5. In the closed inverse subsemigroup M_3, as shown in the Figure below, we have

$$M_3 = M_{3,0} \cup M_{3,1} \cup M_{3,2} \ .$$

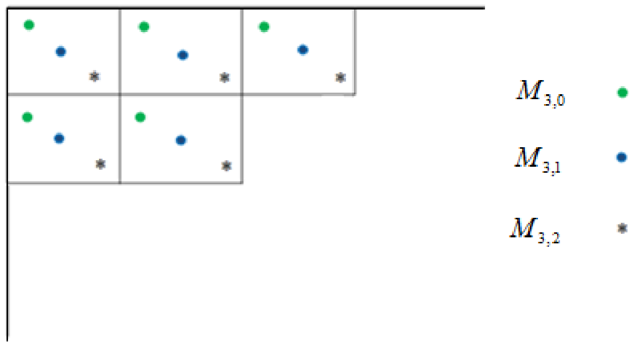

Figure 12: *An insider view of M_3 and its components*

From Theorem 3.4.2, which tells us that M_d is a disjoint union of d copies of B, it is pretty straightforward to see that

$$M_3 \cong B \cup B \cup B.$$

Therefore M_3 is isomorphic exactly to three copies of the bicyclic monoid.

We take some examples of different M_d's to find their generating sets.

Example 3.4.6. Every element of M_2 can be written as a product of

$$x = (0,2),\ y = (1,3) \quad \text{and} \quad x^{-1} = (2,0),\ y^{-1} = (3,1).$$

So, x and y generate M_2, as shown in the following square grid.

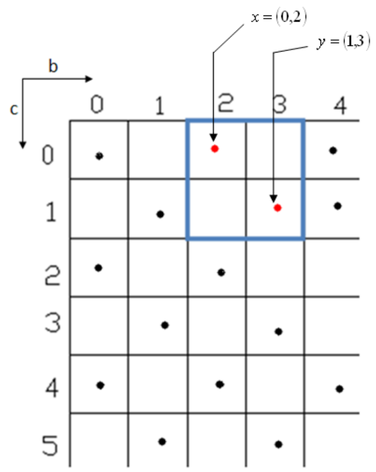

Figure 13: *The generating set of the inverse subsemigroup M_2*

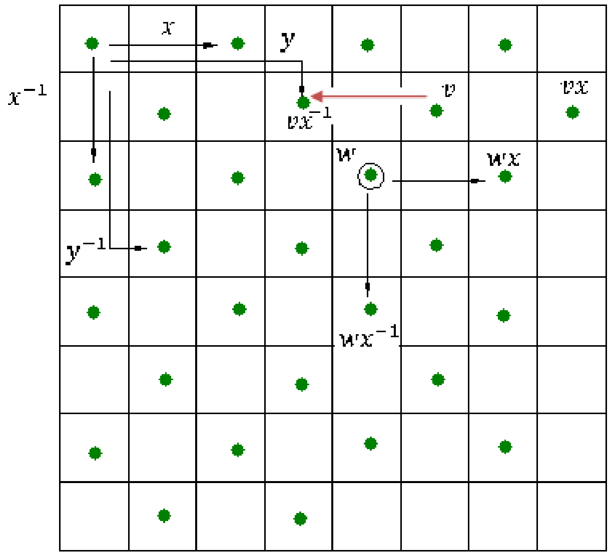

Figure 14: *Moving around in the grid defining* M_2

Example 3.4.7. In the case of M_4, the generating set consists of the elements

$$x = (0, 4),\ y = (1, 5),\ z = (2, 6),\ w = (3, 7).$$

and their inverses. This idea is illustrated in Figure 15.

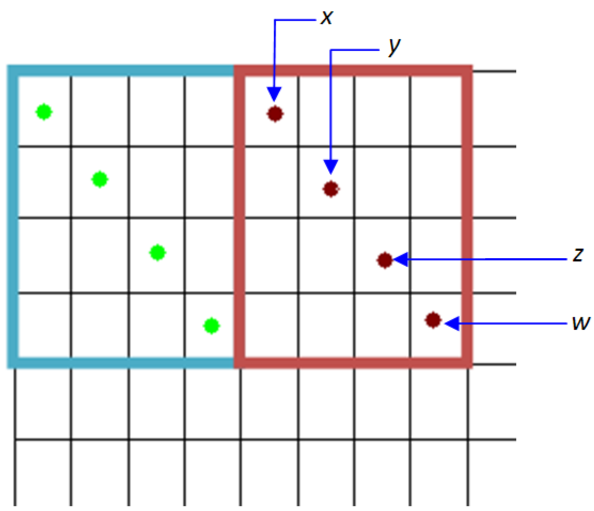

Figure 15: *The generating set of the inverse subsemigroup* M_4

Next, we see the interaction between the elements of distinct copies of B in M_d.

Remark 3.4.8. We will see how elements in different copies of B multiply together. Label the copies $B_0, B_1, \ldots, B_{d-1}$. Take $(a + rd, a + sd) \in B_a$, $(b + td, b + ud) \in B_b$. There are two cases to be considered:

(i) If $a + sd \geqslant b + td$, then

$$\begin{aligned}(a + rd, a + sd) \,.\, (b + td, b + ud) &= (a + rd, a + sd - b - td + b + ud)\\ &\quad \text{(by the multiplication rule on } B)\\ &= (a + rd, a + (s - t + u)d) \in B_a\\ &\quad (\text{since } s - t + u \geqslant 0)\ .\end{aligned}$$

(ii) If $a + sd < b + td$, then

$$\begin{aligned}(a + rd, a + sd) \,.\, (b + td, b + ud) &= (a + rd + b + td - a - sd, b + ud)\\ &\quad \text{(from the multiplication rule in } B)\\ &= (b + (r + t - s)d, b + ud) \in B_b\\ &\quad (\text{since } r + t - s \geqslant 0)\ .\end{aligned}$$

It follows that the product of these two elements should strictly belong to one of B_a or B_b.

3.5 Bruck-Reilly Semigroups

We now seek to extend our previous results into a broader class, namely the Bruck-Reilly semigroups. A Bruck-Reilly semigroup $\mathrm{BR}(G, \theta)$ is constructed by fixing a group G and a homomorphism $\theta : G \to G$. The elements are of the form (m, g, n) where $m, n \in \mathbb{N}$ and $g \in G$.

The multiplication rule for these forms is as follows:

$$(m,g,n)(r,h,s)=\begin{cases}(m,g(h\theta^{n-r}),n-r+s) & \text{if} \quad n\geqslant r\,,\\ (m+r-n,(g\theta^{r-n})h,s) & \text{if} \quad n\leqslant r\,.\end{cases}$$

It is easy to see that the idempotents in $\mathrm{BR}(G,\theta)$ are the elements of the form (k,e_G,k), where e_G is the identity element of G.

In the case of $\theta = id_G$, (the identity map on G) the multiplication simplifies to

$$(m,g,n)\,(r,h,s)=\begin{cases}(m,gh,n-r+s) & \textit{if} \quad n\geqslant r\,,\\ (m+r-n,gh,s) & \textit{if} \quad n\leqslant r\,.\end{cases}$$

Hence in this case,

$$\mathrm{BR}(G,id_G)\cong\ B\times G,$$

where an isomorphism is given by

$$(m,g,n)\ \longmapsto((m,n),g).$$

Another fairly simple case of $\mathrm{BR}(G,\theta)$ is when $\theta=0$, that is $g\theta=e_G$ for all $g\in G$. Then we have

$$(m,g,n)\,(r,h,s)=\begin{cases}(m,g,n-r+s) & \textit{if} \quad n>r\,;\\ (m,gh,s) & \textit{if} \quad n=r\,;\\ (m+r-n,h,s) & \textit{if} \quad n<r\ .\end{cases}$$

We can extend the diagrammatic representation of B used earlier in this chapter, by representing the element $(m,g,n)\in\mathrm{BR}(G,\theta)$ by a cell in an infinite square grid in row m and column n containing the element g.

Lemma 3.5.1. *In a Bruck-Reilly semigroup* $\mathrm{BR}(G,\theta)$ *the natural partial order is given by*

$$(m,g,n) \leqslant (p,h,q) \iff m \geqslant p,\ m-n = p-q \textit{ and } g = h\,\theta^{m-p}\,.$$

Proof. We have $(m,g,n) \leqslant (p,h,q)$ if and only if $(m,g,n) = (k,e_G,k)(p,h,q)$, for some k. But

$$(k,e_G,k)(p,h,q) = \begin{cases} (k, h\theta^{k-p}, k-p+q) & \text{if} \quad k \geqslant p\,, \\ (k+p-k, (e_G\,\theta^{p-k})h, q) & \text{if} \quad k \leqslant p\,. \end{cases}$$

In the second case, we can only have $(m,g,n) = (p,(e_G\,\theta^{p-k})h,q)$ when $p = k$, and this is contained in the first case, in which

$$(m,g,n) = (k, h\theta^{k-p}, k-p+q)\,.$$

Hence $m = k \geqslant p$, $m-p+q = n$ and $g = h\theta^{m-p}$, as we claimed. $\square$

In the representation of elements of $\mathrm{BR}(G,\theta)$ in a grid, to go downwards in the natural partial order we move diagonally down and to the right, and apply θ for each row descended.

Proposition 3.5.2. *The subset* $B = \{(m,e_G,n) : m,n \in \mathbb{N}\}$ *of* $\mathrm{BR}(G,\theta)$ *is an inverse subsemigroup isomorphic to the bicyclic monoid* B*, and* B *is a closed inverse subsemigroup if and only if* θ *is injective.*

Proof. It is clear from the form of the multiplication in $\mathrm{BR}(G,\theta)$ that B is an inverse subsemigroup isomorphic to B. From Lemma 3.5.1, B is closed provided that $h\theta^k = e_G$ if and only if $h = e_G$, and this occurs if and only if θ is injective. $\square$

We will proceed to examine the structure of subsemigroups in different versions of $\mathrm{BR}(G,\theta)$ in greater detail. And then we will see that understanding these structures is the key to construct their cosets and count the index.

Theorem 3.5.3. *If G is finite, and θ is injective, then the copy of the bicyclic monoid B in* $\mathrm{BR}(G, \theta)$ *has index* $|G|$.

Proof. We aim to describe the cosets of the copy of the bicyclic monoid

$$B = \{(m, e_G, n) : m, n \geqslant 0\}$$

in $\mathrm{BR}(G, \theta)$. The assumptions that $\theta : G \to G$ is injective and that G is finite then imply that θ is also surjective. Therefore θ is an automorphism of G. Proposition 3.5.2 ensures that B is closed.

We will begin by constructing the coset of B determined by an element (p, g, q) which will be of the form $(B(p, g, q))^{\uparrow}$ and p, g, q are now fixed.

$$\begin{aligned} (m, e_G, n)(p, g, q) &= (m, g\theta^{n-p}, q + n - p) \text{ if } n \geqslant p \\ &= (m - n + p, g, q) \text{ if } n \leqslant p\,. \end{aligned}$$

The situation here can be viewed through the following Figure.

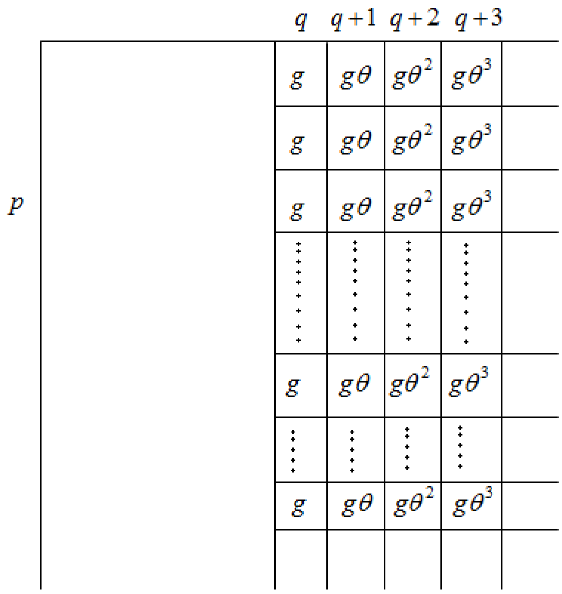

Figure 16: *The cell-elements of the coset* $(B(p, g, q))^{\uparrow}$

It implies that $(B(p,g,q))^{\uparrow}$ contains every cell at every level, because the upwards closure has been taken here. And now the next step is to determine the elements of G.

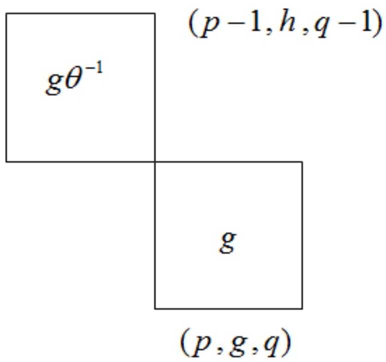

Figure 17: *Two different cell-elements in two different levels of the coset* $(B(p,g,q))^{\uparrow}$

We observe form the simple sketch above that the elements of G will be of these forms:
every cell contains $g\,\theta^{-1}$ in column $(q-1)$,
every cell contains $g\,\theta^{-2}$ in column $(q-2)$,
$\vdots$
every cell contains $g\,\theta^{-q}$ in column 0,
every cell contains g in column q,
every cell contains $g\,\theta$ in column $(q+1)$,
every cell contains $g\,\theta^{2}$ in column $(q+2)$, and so on.

Therefore every column contains the same element of G in each cell. If column i contains element x then column $i+1$ contains $x\theta$ and so the element g in the column q determines the entire coset. We conclude that there will be one coset for each element of G. It follows that the index of B in $\mathrm{BR}(G,\theta)$ is $|G|$. $\square$

Proposition 3.5.4. *If G is an infinite group, and θ is injective, then the copy B of the bicyclic monoid B in $\mathrm{BR}(G,\theta)$ has infinite index.*

Proof. Here θ need not to be an isomorphism, but as in the finite case, by Proposition 3.5.2, B is closed since θ is injective.

We can repeat the construction of the cosets as in the proof of Theorem 3.5.3. Since B is full, every element determines a coset, and for $g \in G$ the coset $(B(0,g,0))^{\uparrow}$ is the upwards closure of

$$\begin{aligned} B(0,g,0) &= \{(m, e_G, n)(0,g,0) : m, n \geqslant 0\} \\ &= \{(m, g\theta^n, n) : m, n \geqslant 0\}. \end{aligned}$$

But by Lemma 3.5.1, we see that $B(0,g,0)$ is a closed subset, and so

$$(B(0,g,0))^{\uparrow} = \{(m, g\theta^n, n) : m, n \geqslant 0\}.$$

Moreover, for $g \neq h$, the cosets $(B(0,g,0))^{\uparrow}$ and $(B(0,h,0))^{\uparrow}$ are distinct, and therefore there are infinitely many of them. $\square$

Remark 3.5.5. An important difference now arises between the cases of finite and infinite G in $\mathrm{BR}(G,\theta)$. In order that B is a closed inverse subsemigroup, we must have θ injective and then if G is finite as in Theorem 3.5.3, θ must be an isomorphism. However, if G is infinite, θ need not be an isomorphism – that is, θ need not be surjective. In this case, we can construct extra cosets not given by the recipe in the proof of Proposition 3.5.4. Suppose that $g \in G$ and that g is not in the image of θ. Consider the coset $(B(p,g,q))^{\uparrow}$. As before, this is the closure of the set

$$\{(m, g\theta^{n-p}, q+(n-p)) : m, n \in \mathbb{N}, n \geqslant p\} \cup \{(m+p-n, g, q) : m, n, \in \mathbb{N}, n \leqslant p\}.$$

The elements $(m+p-n, g, q)$ form a complete column of cells (in column q) with each cell containing the element g. The elements $(m, g\theta^{n-p}, q+(n-p))$ form complete columns, in columns $q+r$ for $r \geqslant 0$, with each cell in the column containing the element $g\theta^r$.

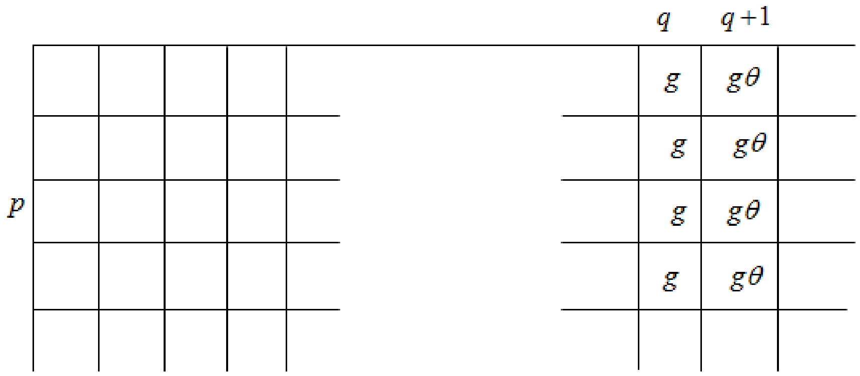

Figure 18

Forming the closure of this set would involve adding cells diagonally upwards and to the left: cells in column $q-1$ would have to contain an element h with $h\theta = g$. By assumption no such h exists and the elements (p, g, q) in column q are maximal. Hence the coset $(B(p, g, q))^{\uparrow}$ is equal to the set

$$\{(m, g\theta^{n-p}, q + (n-p)) : m, n \in \mathbb{N}, n \geqslant p\} \cup \{(m + p - n, g, q) : m, n, \in \mathbb{N}, n \leqslant p\} .$$

and does not contain cells in columns $0, 1, \ldots, q-1$.

Chapter 4

Polycyclic Monoids and Graph Inverse Semigroups

In this chapter, we begin by focusing on the theory of the polycyclic monoids and then on graph inverse semigroups. Some of the results being mentioned are well-known. We re-prove Lawson's characterization of the closed inverse submonoids of the polycyclic monoids [16, Theorem 4.3]. This allows us to investigate the index properties of these inverse submonoids. The characterization of the closed inverse submonoids of the polycyclic monoid has been extended to closed inverse subsemigroups of the graph inverse semigroups. The polycyclic case is dealt with first, as the graph inverse semigroup case is more technical. Some facts and basic definitions about the polycyclic monoids and graph inverse semigroups will be needed in this chapter are introduced in the next few sections.

4.1 Closed Inverse Submonoids of the Polycyclic Monoids

The goal of this section is to describe all the closed inverse submonoids of the polycyclic monoids. Moreover, we determine the index of each type of these closed inverse submonoids whether it is finite or infinite. The next definitions are particularly important in order to understand what follows in this chapter.

Definition Let A^* be the free monoid on the alphabet A and suppose that $w \in A^*$ with

$w = uv \in A^*$. Then u is called a *prefix* of w and v is called a *suffix* of w. Two words w_1 and w_2 are called *suffix comparable*, if one is a suffix of the other.

Definition Any two closed inverse subsemigroups H and K of an inverse subsemigroup S are called *conjugate,* if there exists $s \in S$ such that

$$sHs^{-1} \subseteq K \text{ and } s^{-1}Ks \subseteq H\,.$$

Definition The *polycyclic monoid* P_n , where $n \geqslant 2$, is defined as a monoid with zero by the presentation:

$$P_n = \langle a_1, \ldots, a_n, a_1^{-1}, \ldots, a_n^{-1} : a_i a_i^{-1} = 1 \text{ and } a_i a_j^{-1} = 0,\, i \neq j \rangle\,.$$

Each non-zero element of P_n has a unique representation as $u^{-1}v$ for $u, v \in A^*$. The bicyclic monoid B studied in Chapter 3 and defined by the presentation

$$B = \langle a_1, a_1^{-1} : a_1 a_1^{-1} = 1 \rangle\,,$$

can be thought of the polycyclic monoid P_1 , but here for $n = 1$ we do not need the zero.

The closed inverse submonoids of P_n were classified by Lawson in [16, Theorem 4.3]. In this section, we describe this classification and then we use it to determine the index of each closed inverse submonoid. We begin by constructing the main ingredients of our approach.

Let $n \geqslant 1$ and let $A = \{a_1, \ldots, a_n\}$ be an alphabet, and let A^* be the free monoid on A. We order A^* using the *suffix ordering* as follows:

$$u \geqslant v \quad \Longleftrightarrow \quad \exists p \in A^* \quad \text{with} \quad v = pu\,.$$

Hence a word is less than all its proper suffixes, or equivalently, adding a prefix to a word moves it down in this partial order. The empty word 1 is the maximum, and the letters in A are the non-empty maximal words.

We think of A^* as an $n-$ary tree rooted at 1: a word $w \in A^*$ has n descendants $a_1w, \ldots, a_nw$ in this tree. A chain in the tree A^* is then a sequence of words (w_n) satisfying $w_0 = 1$ and $w_n = a_{j_n} w_{n-1}$, for all $n \geqslant 1$.

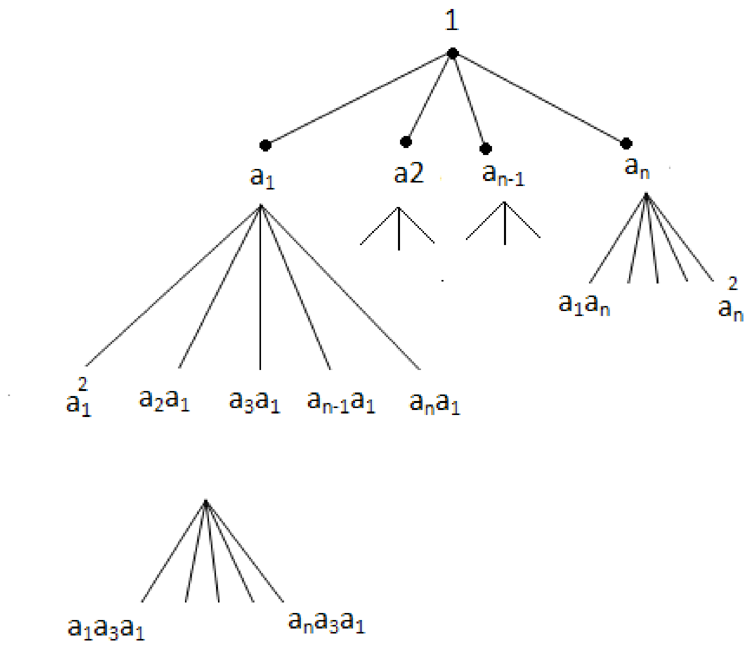

Figure 19: *The tree of idempotents* A^*

Example 4.1.1. In the tree representing A^* when $A = \{a, b, c, d\}$ below, we notice that the word $w = abcabd \leqslant bcabd \leqslant cabd \leqslant abd \leqslant bd \leqslant d \leqslant 1$.

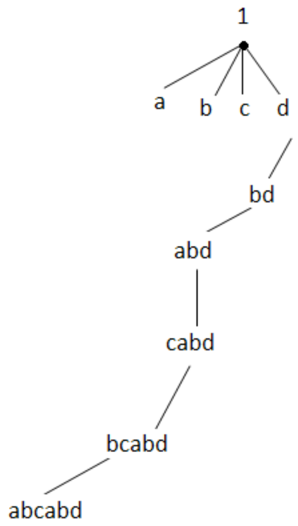

Figure 20: *The branch contains the word $w = abcabd$ on the tree A^**

The non-zero elements of the polycyclic monoid P_n can be represented as strings $u^{-1}v$ with $u, v \in A^*$. For the following discussion it is more convenient to consider such a string as an ordered pair of strings (u, v). Hence

$$P_n = \{(u, v) : \; u, v \in A^*\} \cup \{0\} \, .$$

The non-zero idempotents of P_n are the repeated pairs (u, u) and so we can identify $E(P_n)$ with $A^* \cup \{0\}$, writing u rather than the repeated pair (u, u).

As mentioned in Section 1.10, it is helpful to regard (u, v) as an arrow from u to v with the ordering on arrows given by

$$(u, v) \geqslant (x, y) \quad \iff \quad \exists p \in A^*, \;\; \text{with} \;\; x = pu \;\; \text{and} \;\; y = pv \, ,$$

and with $(u, v) \geqslant 0$ for all $u, v \in A^*$. The ordering on $E(P_n) \setminus \{0\}$ is then the suffix ordering on A^*. An arrow (u, v) from u to v combines with an arrow (v, y) from v to y to give the arrow (u, y) from u to y .

To compose (u, v) and (x, y) we find the greatest lower bound in A^* of v and x. We obtain non-zero lower bounds for v by adding prefixes to v and hence a non-zero lower bound for v and x is a word l of the form $l = pv = qx$, for some $p, q \in A^*$. Hence either v is a suffix of x with $x = p'v$ or x is a suffix of v with $v = q'x.$ If neither of these occurs, then x, v have a greatest lower bound 0 and so $(u, v)(x, y) = 0$. But if $x = p'v,$ we have

$$(u, v)\,(x, y) = (p'u, p'v)\,(x, y) = (p'u, y)\,,$$

since an arrow $(p'u, p'v)$ from $p'u$ to $p'v$ combines with an arrow $(p'v, y)$ from $p'v$ to y to give the arrow $(p'u, y)$ from $p'u$ to y. Likewise, if $v = q'x,$ we have

$$(u, v)\,(x, y) = (u, v)(q'x, q'y) = (u, q'y)\,.$$

This multiplication can, of course, be obtained more directly as the consequence of the defining relations

$$a_i a_i^{-1} = 1\,, \quad a_i a_j^{-1} = 0 \;(i \neq j)\,,$$

for P_n and interpreting the ordered pair (u, v) as the word $u^{-1}v$. Then if $(u^{-1}v)\,(x^{-1}y) \neq 0,$ then one of v, x must exhaust the other in cancellation by following the rule $a_i a_i^{-1} = 1,$ and so one is a suffix of the other, as before.

A closed inverse subsemigroup of P_n must be a *submonoid*, since $1 \in P_n$ is the maximal idempotent. Lawson classified the closed inverse submonoids of P_n in [16, Theorem 4.3].

Theorem 4.1.2. *In the polycyclic monoid, there are exactly three types of proper closed inverse submonoids L:*

(1) Finite chain type: L consists of a finite chain of idempotents.

(2) Infinite chain type: L consists of an infinite chain of idempotents.

(3) Cycle type: there exists $x, p \in A^*$ with no non-trivial common prefix, such that every non-zero element of L is either an idempotent (q, q) for some suffix q of x or has the form $(qp^r x, qp^s x)$ for all $r, s \geqslant 0$, with q a suffix of p. In this case we write $L = L_{p,x}$.

Proof. We give a slightly different proof to that in [16, Theorem 4.3].

It is easy to see that each of the three types described in the theorem is indeed a proper closed inverse submonoid of P_n.

Suppose that $u, v \in E(L)$ and that neither is a suffix of the other: then their product in $E(L)$ is 0 and hence $0 \in E(L)$. Since L is closed, we deduce that $L = P_n$, since all the elements $(u, v) \geqslant 0$ for all $u, v \in A^*$. Hence if L is proper, $L \neq P_n$, and $u, v \in E(L) \subset A^*$, then in A^* we must have one of u, v is a suffix of the other and so either $u \leqslant v$ or $v \leqslant u$ in the suffix order. Therefore, $E(L)$ is a chain and u, v are on the same branch of the tree of idempotents describing $E(L)$.

Thus, if $L = E(L)$, two cases arise:

- L is a finite chain in A^*,
- L is an infinite chain of idempotents.

Therefore, these are the finite and infinite chain types, as Lawson showed in [16].

Now, if $L \neq E(L)$, we think of a non-idempotent element of L as an arrow (u, v) between two vertices of the chain $E(L)$, as in the Figure below.

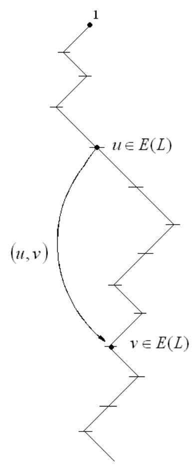

Figure 21: *A non-idempotent element* (u, v) *of* L

Then there exists a non-idempotent element $(x, y) \in L$, and we can assume that y is longer than x in A^*. We observe that $(x, y)\,(x, y)^{-1} = (x, x) \in L$ and $(x, y)^{-1}(x, y) = (y, y) \in L$ and so that $x, y \in A^*$ represent elements of $E(L)$. Hence x, y are on the idempotent chain determined by $E(L)$ and then $y \leqslant x$ in the natural partial order on $E(L)$, so that we have $y = px$ for some $p \in A^*$. Amongst such elements (x, y), choose the element x to have the shortest possible length, and $y = px$ with p of minimal length, so that $(x, px) \in L$, with $p \neq 1$.

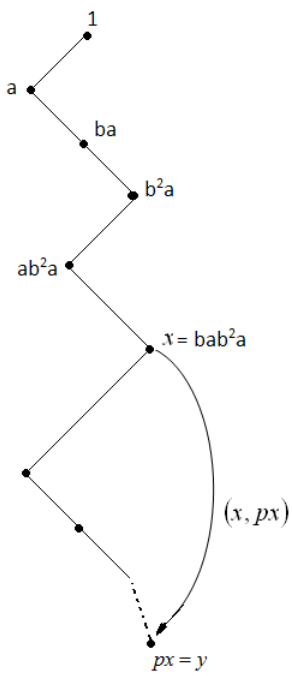

Figure 22: *An example of a choice of the arrow* (x, px)

Then $px \in E(L)$ and so $(px, px)(x, px) = (px, p^2x) \in L,$ and repeating this we deduce that for all $r \geqslant 0$ we have $(p^r x, p^{r+1}x) \in L.$

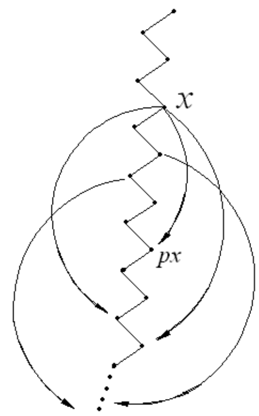

Figure 23: *The process of restricting any two elements on the chain*

Straighten the chain:

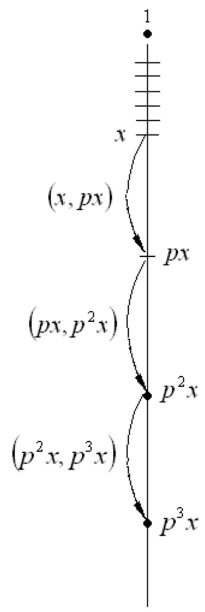

Figure 24: *A chain of some non-idempotent elements of* L

So $(p^r x, p^{r+1}x)\,(p^{r+1}x, p^{r+2}x) = (p^r x, p^{r+2}x) \in L$ and hence by repeating the same process we get $(p^r x, p^s x) \in L$, for all $r, s \geqslant 0$.

Since $p^s x \in E(L)$, it follows that $E(L)$ contains words of unbounded length and so $E(L)$ is of infinite chain type. In other words, $E(L)$ is represented by all words on an infinite chain in A^*. In particular, $E(L)$ contains each suffix of x and each suffix of $p^s x$ for any $s \geqslant 0$ and so each word of the form $qp^k x$, where $k \geqslant 0$, and q is a suffix of p. Then L contains all elements of the form $(qp^r x, qp^s x)$ for all $r, s \geqslant 0$, since

$$(qp^r x, qp^r x)\,(p^r x, p^s x) = (qp^r x, qp^s x)$$

as Figure 25 and 26 show.

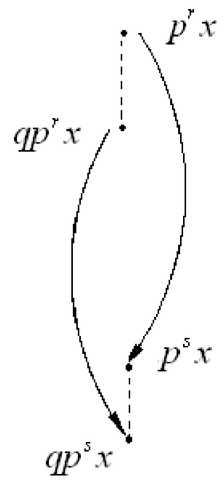

Figure 25: $qp^r x$ *restricting* $(p^r x, p^s x)$

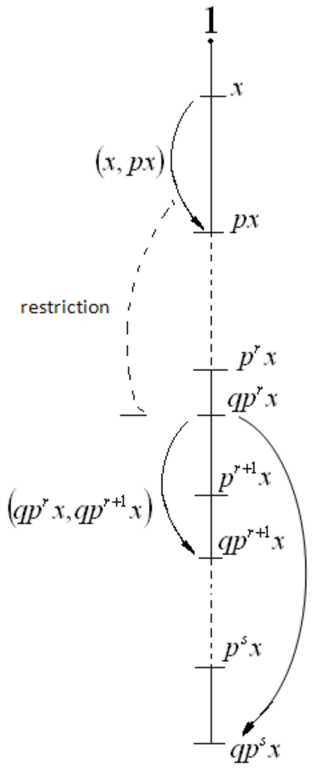

Figure 26: $qp^r x$ *restricting* (x, px)

It then follows that $L_{p,x} \subseteq L$.

If $(u, v) \in L$ with $|u| < |x|$ then by the choice of x, we have $(u, v) \in E(L)$, so $u = v$ and u is a suffix of x. If $|u|, |v| \geqslant |x|$ then u, v are suffix comparable with $p^k x$ for all $k \geqslant 0$. Choosing large enough k, we have u, v are suffixes of $p^k x$ and so $u = q_1 p^r x$ and $v = q_2 p^s x$ for some $r, s \geqslant 0$ and suffixes q_1, q_2 of p. Suppose that $|q_2| > |q_1|$ and write $p = yq_2$. Then

$$(q_1 x, q_1 p^r x)\,(q_1 p^r x, q_2 p^s x)\,(q_2 p^s x, q_2 x) = (q_1 x, q_2 x) \in L$$

and so $(x, px)\,(q_2 x, q_1 x) = (x, yq_2 x)\,(q_2 x, q_1 x) = (x, yq_1 x) \in L$.

But $|yq_1| < |yq_2| = |p|$, and so this contradicts the choice of p. Hence $|q_1| = |q_2|$ and since they are suffix comparable, $q_1 = q_2$. Therefore $(u, v) \in L_{p,x}$ and we conclude that $L = L_{p,x}$. □

We now consider the index of each of the types of closed inverse submonoids of P_n.

Theorem 4.1.3. *A closed inverse submonoid L of the polycyclic monoid P_n of finite chain type has infinite index.*

Proof. A closed inverse submonoid L of finite chain type is determined by a word $w \in A^*$ and then L has the form

$$L = \{(u,u) : u \text{ is a suffix of } w\}.$$

A possible coset representative for L has the form (u,v) where u is a suffix of w, but there is no restriction on v, it can be an arbitrary element of A^*. Consider the set

$$T(L) = \{(u,v) : u \text{ is a suffix of } w \text{ and } v \in A^*\},$$

which contains all possible representatives of cosets of L in P_n.

Choose $(u_1,v_1),(u_2,v_2) \in T(L)$, then

$$L(u_1,v_1)^{\uparrow} = L(u_2,v_2)^{\uparrow} \quad \Longleftrightarrow \quad (u_1,v_1)(u_2,v_2)^{-1} = (u_1,v_1)(v_2,u_2) \in L.$$

So if $L(u_1,v_1)^{\uparrow} = L(u_2,v_2)^{\uparrow}$ then v_1, v_2 are suffix comparable. We may suppose that $v_2 = hv_1$ for some $h \in A^*$. Then

$$\begin{aligned}(u_1,v_1)(u_2,v_2)^{-1} &= (u_1,v_1)(hv_1,u_2)\\ &= (hu_1,u_2) \in L\\ &\Longleftrightarrow u_2 = hu_1,\end{aligned}$$

and both u_1, u_2 are suffixes of w. Hence the cosets $L(u_1,v_1)^{\uparrow}$ and $L(u_2,v_2)^{\uparrow}$ coincide if and only if $(u_2,v_2) = (hu_1,hv_1)$ where both u_1 and hu_1 are suffixes of w. Hence for a fixed choice of $(u,v) \in T(L)$, only finitely many other elements of $T(L)$ can represent the same coset and there are infinitely many distinct cosets. $\square$

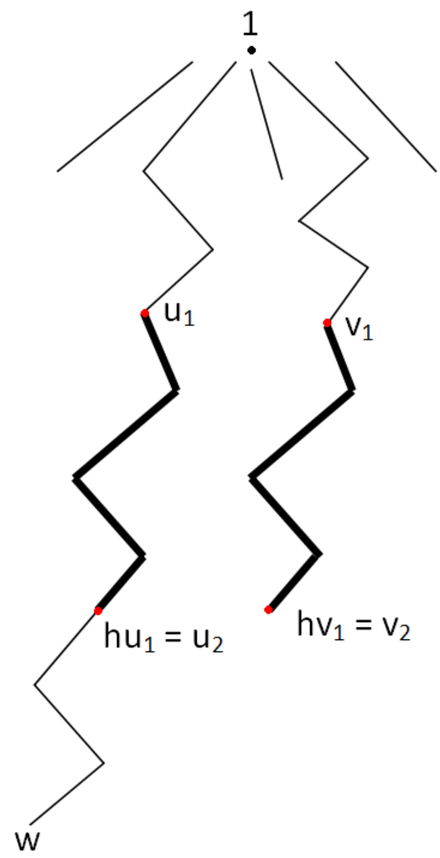

Figure 27:

Theorem 4.1.4. *A closed inverse submonoid L of the polycyclic monoid P_n of infinite chain type has infinite index.*

Proof. Let L be a closed inverse submonoid of infinite chain type in P_n. Then L will be of the form

$$L = \{(u, u) : u \in S\}$$

for some subset $S \subset A^*$ closed under taking suffixes. Hence S forms an infinite chain in the tree representation of A^*. Consider the set $T(L) = \{(w, v) : w \in S, v \in A^*\}$, which contains all possible representatives of cosets of L in P_n.

Choose $(w_1, v_1), (w_2, v_2) \in T(L)$, then

$$L(w_1, v_1)^{\uparrow} = L(w_2, v_2)^{\uparrow} \iff (w_1, v_1)(w_2, v_2)^{-1} = (w_1, v_1)(v_2, w_2) \in L.$$

So v_1, v_2 are suffix comparable, and we may suppose that $v_1 = hv_2$. Then

$$(w_1, v_1)(v_2, w_2) = (w_1, hv_2)(v_2, w_2) = (w_1, hw_2) \in L$$

if and only if $w_1 = hw_2$. Hence the cosets $L(w_1, v_1)^{\uparrow}$ and $L(w_2, v_2)^{\uparrow}$ coincide if and only if $(hw_2, hv_2) = (w_1, v_1)$ where both hw_2 and w_2 are in S. Hence there are infinitely many distinct cosets. □

Theorem 4.1.5. *A closed inverse submonoid $L_{p,x}$ of the polycyclic monoid P_n of cycle type has infinite index.*

Proof. Let

$$L_{p,x} = \{\, (qp^r x, qp^s x) : \ r, s \geqslant 0 \text{ with } q \text{ a suffix of } p\} \cup \{(q, q) : q \text{ a suffix of } x\}.$$

be a closed inverse subsemigroup of the cycle type. Suppose that p does not contain the generator $a_m \in A$. Consider the cosets

$$C_k = L_{p,x}\, (px, a_m^k)^{\uparrow}\, .$$

Let $k > l$. Now $C_k = C_l$ if and only if

$$(px, a_m^k)\, (px, a_m^l)^{-1} = (px, a_m^k)\, (a_m^l, px) = (px, a_m^{k-l} px) \in L$$

Now if $(px, a_m^{k-l} px) \in L$, then

$$a_m^{k-l} = p^t, \tag{4.1.1}$$

for some t. If we assume that the word p does not include a_m, then this implies that equation (4.1.1) is impossible, and so

$$C_k \neq C_l\ .$$

Then the cosets $C_k = L_{p,x}\,(px, a_m^k)^{\uparrow}$ are all distinct and so $L_{p,x}$ has infinite index.

Now suppose that p uses all the generators in A. We choose a word $w \in A^*$ and define

$$C_k = L_{p,x}\,(px, w^k)^{\uparrow}\,.$$

Repeating the argument above with w replacing a_m we deduce that

$$C_k = C_l \quad\Longleftrightarrow\quad w^{k-l} = p^t\,.$$

Hence, if we choose a w that is not a subword of p^t we guarantee that $C_k \neq C_l$. Since the number of subwords of p^t is finite, we can certainly choose such a w. Therefore, this shows that we construct an infinite list of cosets $C_k = L_{p,x}\,(px, w^k)^{\uparrow}$ that are all distinct and so $L_{p,x}$ has infinite index. Consequently, a cycle type closed inverse subsemigroup of P_n has infinite index. $\square$

4.2 Closed Inverse Subsemigroups of Graph Inverse Semigroups

Graph inverse semigroups were first defined by Ash and Hall [2]. They generalize the polycyclic monoids P_n for $n \geqslant 2$, which are the graph inverse semigroups obtained from one-vertex directed graphs.

Definition A directed graph Γ consists of a set $X = \mathrm{vert}(\Gamma)$ of *vertices* and a set $Y = \mathrm{edge}(\Gamma)$ of *edges*, and two maps $o, t : Y \to X$. The vertex $o(y)$ is called the *origin* of y and the vertex $t(y)$ is called the *terminus* of y. These two vertices are called the extremities of y. We say that the two vertices are *adjacent* if they are the extremities of some edge.

A directed graph is usually represented by a diagram in which an edge $y \in \mathrm{edge}(\Gamma)$ is represented by an arrow from $o(y)$ to $t(y)$.

Definition A *directed path* in Γ is either an element of $\mathrm{vert}(\Gamma)$ (these are the paths of *length* zero) or is a finite sequence of edge $(y_1, y_2, \ldots, y_m)$ such that $t(y_i) = o(t_{i+1})$ for all $i\,, 1 \leqslant i < m$. Such a directed path has *length* m. A *directed circuit* is a path of length $m > 0$ such that $o(y_1) = t(y_m)$.

Definition Given a directed path $p = (y_1, y_2, \ldots, y_m)$ in Γ, a *suffix* of p is a path $s = (y_k, \ldots, y_m)$ for some $k \geqslant 1$, or else the length zero path $t(y_m)$. Two paths p, q are *suffix comparable* if one is a suffix of the other.

Definition Let Γ be a finite directed graph. The graph inverse semigroup $G(\Gamma)$ is defined as follows: as a set we have

$$G(\Gamma) = \{(p, q) : p, q \text{ are paths in } \Gamma \text{ with the same origin}\} \cup \{0\}\,.$$

and the multiplication on $G(\Gamma)$ is given by

$$(m, n)\,(p, q) = \begin{cases} 0 & \text{if } n, p \text{ are not suffix comparable,} \\ (m, n'q) & \text{if} \quad n = n'p\,; \\ (p'm, q) & \text{if} \quad p = p'n\ . \end{cases}$$

It is easy to check that $G(\Gamma)$ is indeed an inverse semigroup in which $(p, q)^{-1} = (q, p)$ and the idempotents are the pairs (p, p).

The multiplication in $G(\Gamma)$ is illustrated below.

An element $(u, v) \in G(\Gamma)$ is represented by a pair of labelled arrows with the same starting point.

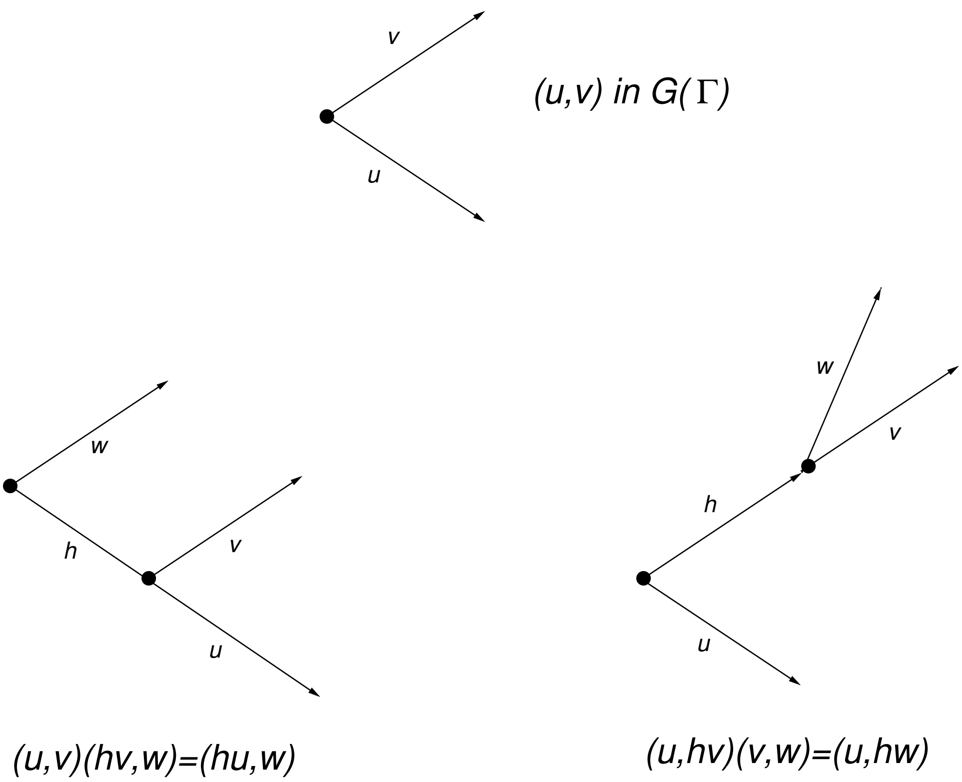

Figure 28: *Multiplication in* $G(\Gamma)$

We note that if $(u,u),(v,v) \in G(\Gamma)$ then $(u,u)(v,v) = 0$ unless u, v are suffix comparable. Then if $v = hu$ we have $(u,u)(v,v) = (v,v)$.

For the natural partial order we have $(m,n) \geqslant (p,q)$ if and only if p, q share a common prefix y such that $ym = p$ and $yn = q$. (That is, we move up in the partial order by deleting common prefixes.) It follows that graph inverse semigroups are E^*–unitary (see Section 1.3). The induced ordering on the set of directed paths in Γ is again the suffix ordering: $p \leqslant m$ if and only m is a suffix of p.

Our first result generalizes Lawson's classification [16, Theorem 4.3] of closed inverse submonoids of the polycyclic monoids P_n to the closed inverse subsemigroups of graph inverse semigroups $G(\Gamma)$.

Theorem 4.2.1. *In a graph inverse semigroup $G(\Gamma)$ there are three types of proper closed inverse subsemigroups L:*

(1) Finite chain type: L consists of a finite chain of idempotents.

(2) Infinite chain type: L consists of an infinite chain of idempotents.

(3) Cycle type: L has the form

$$L = L_{p,d} = \{\, (vp^r d, vp^s d) : r, s \geqslant 0 \text{ with } v \text{ a suffix of } p\,\} \cup \{(q, q) : q \text{ a suffix of } d\},$$

where p is a directed circuit in Γ, d is a directed path in Γ starting at the initial point of p, and where p, d don't share a non-trivial prefix.

Proof. The proof is similar to that given for Theorem 4.1.2. As in that proof, it is easy to see that each of the three types described in this theorem is indeed a proper closed inverse submonoid of $G(\Gamma)$.

Let L be a closed inverse subsemigroup of $G(\Gamma)$. If w and w' are paths occurring in elements of L and are not suffix comparable, then the product of the idempotents (w, w) and (w', w') in L is equal to 0, and so $0 \in L$ and by closure $L = G(\Gamma)$. Hence we may assume that any two paths occurring in elements of L are suffix comparable and hence have the same terminal vertex. By definition, if $(u, v) \in G(\Gamma)$ then u, v have the same initial vertex. Hence we may assume that if $(u, v) \in L$ then u, v have the same initial and the same terminal vertex in Γ.

Suffix comparability then ensures that any proper closed inverse subsemigroup of $G(\Gamma)$ consisting entirely of idempotents is either a finite or an infinite chain. In the second case, in order to obtain directed paths of arbitrary length, Γ must contain a directed cycle.

We shall now describe those closed inverse subsemigroups of $G(\Gamma)$ which contain non-idempotent elements. Suppose that $L \neq E(L)$ is a closed inverse subsemigroup of $G(\Gamma)$. So there exists $(u, v) \in L$, with $u \neq v$, and we shall assume that the path u is shorter than the path v. Hence u is a suffix of v and so $v = pu$ for some path p. Since u and v have the same initial and terminal vertices, p must be a directed cycle in Γ. If p and u share a common prefix, with $p = ap_1$ and $u = au_1$ then

$$(u_1, p_1u) \geqslant (au_1, ap_1u) = (u, pu)$$

and so by closure, $(u_1, p_1u) \in L$.

Amongst the non-idempotent elements $(u, pu) \in L$, choose $u = d$ to have smallest possible length, and then having chosen d, choose p to be a non-empty directed circuit of smallest possible length. Then d, p do not share a non-trivial prefix. Now for any $m \geqslant 0$ we have $(p^m d, p^m d) \in E(L)$ and so, if $(w_1, w_2) \in L$, each w_i is a suffix of some directed path $p^{m_i} d$. Since L is closed, every suffix of d is in L and by minimality of $|d|$, every element of L that contains a suffix of d is an idempotent (q, q). Hence if $|w_i| < |d|$ we have $w_1 = w_2$. So we may now assume that for $i = 1, 2$ we have $|w_i| \geqslant |d|$, and so $w_1 = up^r d$, $w_2 = vp^s d$ for some $r, s \geqslant 0$ and suffixes u, v of p. $\square$

From the result of the previous theorem, we may immediately conclude the following:

Corollary 4.2.2. *If the graph Γ contains no directed cycle, then every proper closed inverse subsemigroup of $G(\Gamma)$ is a chain of idempotents.*

We have proved the following theorem, based on [16, Theorem 4.4] which treats the polycyclic monoids, to classify the closed inverse subsemigroups of a graph inverse semigroup up to conjugacy. We begin with the following definitions

Definition Let $L = (u, u)^{\uparrow}$ be a closed inverse subsemigroup of finite chain type in a graph inverse semigroup $G(\Gamma)$. We call the initial vertex of the directed path u the *root* of L.

Definition Two directed circuits p, q in Γ are *conjugate* if $p = uv$ then $q = vu$.

Lemma 4.2.3. *Let S be an E^*–unitary inverse semigroup. If H and K are conjugate closed inverse subsemigroups of S with $H \neq S \neq K$ and $H \subseteq E(S)$ then $K \subseteq E(S)$. Moreover, if H has a minimum idempotent, then so does K.*

Proof. There exists $s \in S$ with $s^{-1}Hs \subseteq K$ and $sKs^{-1} \subseteq H$. Let $k \in K$: then $k \neq 0$ and $sks^{-1} \in H$ and so $sks^{-1} \in E(S)$. It follows that $s^{-1}(sks^{-1})s = (s^{-1}s)k(s^{-1}s) \in E(S)$ and $(s^{-1}s)k(s^{-1}s) \leqslant k$. Since S is E^*–unitary, we deduce that $k \in E(S)$.

Suppose that $m \in H \subseteq E(S)$ is the minimum idempotent and that $e \in K$. Then $m \leqslant ses^{-1}$ and so

$$s^{-1}ms \leqslant s^{-1}ses^{-1}s = es^{-1}s \leqslant e$$

and so $s^{-1}ms$ is a minimum idempotent in K. $\square$

Theorem 4.2.4. *(1) Let L be a closed inverse subsemigroup of $G(\Gamma)$ of finite chain type. Then all closed inverse subsemigroups conjugate to it are of finite chain type. Two closed inverse subsemigroups $L = (u,u)^{\uparrow}$ and $K = (v,v)^{\uparrow}$ are conjugate in $G(\Gamma)$ if and only if they have the same root.*

(2) Let L be a closed inverse subsemigroup of $G(\Gamma)$ of infinite chain type. Then only closed inverse subsemigroups conjugate to it are also of infinite chain type. Two closed inverse subsemigroups of infinite chain type are conjugate if and only if there are idempotents $(s,s) \in L$ and $(t,t) \in K$ such that for all paths p in Γ, we have that $(ps,ps) \in L$ if and only if $(pt,pt) \in K$.

(3) Let L be a closed inverse subsemigroup of $G(\Gamma)$ of cycle type. The only closed inverse subsemigroups conjugate to it are of cycle type. Moreover, $L_{p,d}$ is conjugate to $L_{q,k}$ if and only if p and q are conjugate directed circuits in Γ.

Proof. (1) It follows from Lemma 4.2.3 that if L has finite change type then so does every conjugate closed inverse subsemigroup .

Suppose that L and K have the same root $x_0 \in V(\Gamma)$. Then $(u,v) \in G(\Gamma)$, and for any suffix w of u we have

$$(v,u)(w,w)(u,v) = (v,v) \in K\,.$$

Similarly, for any suffix t of v, $(u,v)(t,t)(v,u) = (u,u) \in L$. Hence L and K are conjugate.

Conversely, suppose that L and K are conjugate, with conjugating element $(p,q) \in G(\Gamma)$, so that for any suffixes w of u and t of v we have

$$(q,p)(w,w)(p,q) \in K \text{ and } (p,q)(t,t)(q,p) \in L\,.$$

Let y_0 be the end veretx of u. Then $(q,p)(y_0,y_0)(p,q) \in K$, so that p and y_0 are suffix-comparable: hence p also ends at y_0, and $(q,p)(y_0,y_0)(p,q) = (q,q) \in K$. Therefore q is a suffix of v. Similarly, p is a suffix of u.

Let $v = v_1q$: then

$$(p,q)(v,v)(q,p) = (p,q)(v_1q, v_1q)(q,p) = (v_1p, v_1p) \in L$$

and so v_1p is a suffix of u. Let $u = u_0v_1p$: then

$$(q,p)(u,u)(p,q) = (q,p)(u_0v_1p, u_0v_1p)(p,q) = (u_0v_1q, u_0v_1q) \in K$$

and so u_0v_1q is a suffix of v. But $v = v_1q$ and so u_0 is a vertex (namely the root of L), and $u = v_1p$. Hence u and v have the same initial vertex, and so L and K have the same root.

(2) By Lemma 4.2.3 any closed inverse subsemigroup K conjugate to L must be of chain type, and by part (1) K must be infinite. Suppose that $(t,s)L(s,t) \subseteq K$ and $(s,t)K(t,s) \subseteq L$. Since $0 \notin L$ we have, for all $(u,u) \in L$, that s is suffix comparable with u and similarly for all $(v,v) \in K$, that t is suffix comparable with v. If we consider u with $|u| \geqslant |s|$ then s must be a suffix of u and by closure of L we have $(s,s) \in L$. Similarly $(t,t) \in K$. Suppose that $(ps,ps) \in L$. Then $(t,s)(ps,ps)(s,t) = (pt,pt) \in K$ and similarly if $(pt,pt) \in K$ then $(ps,ps) \in L$.

Conversely, if s and t exist as in the theorem and $(w,w) \in L$ then s is suffix comparable with w.

If w is a suffix of s, with $s = hw$, then

$$(t,s)(w,w)(s,t) = (t,hw)(w,w)(hw,t) = (t,t) \in K$$

and if s is a suffix of w with $w = ps$ then $(ps,ps) \in L$ and so

$$(t,s)(w,w)(s,t) = (t,s)(ps,ps)(s,t) = (pt,pt) \in K\,.$$

Similarly $(s,t)K(t,s) \subseteq L$, and L and K are conjugate.

(3) Suppose that the closed inverse subsemigroups $L_{p,d}$ and $L_{q,k}$ are conjugate in $G(\Gamma)$, and so there exists $(s,t) \in G(\Gamma)$ such that

$$(t,s)L_{p,d}(s,t) \subseteq L_{q,k} \tag{4.2.1}$$

$$(s,t)L_{q,k}(t,s) \subseteq L_{p,d}. \tag{4.2.2}$$

Since $L_{q,k}$ is closed and $L_{p,d}$ is the smallest closed inverse subsemigroup of $G(\Gamma)$ containing (d,pd), then (4.2.1) is equivalent to $(t,s)\,(d,pd)\,(s,t) \in L_{q,k}$. Also, since $0 \notin L_{q,k}$ we must have s suffix-comparable with u and v whenever (u,v) is an element of $L_{p,d}$. Hence $(s,s) \in L_{p,d}$, and similarly $(t,t) \in L_{q,k}$.

First suppose that $s = up^a d$ and $t = vq^b k$ for some $a,b \geqslant 0$, where u is a suffix of p and v is a suffix of q. Write $p = hu$: then

$$\begin{aligned}(t,s)\,(d,pd)\,(s,t) &= (vq^bk, up^ad)\,(d,pd)\,(up^ad, vq^bk)\\ &= (vq^bk, up)\,(u, vq^bk)\\ &= (vq^bk, uhvq^bk) \in L_{q,k}.\end{aligned}$$

It follows that $uhvq^bk = vq^mk$ for some $m \geqslant 0$. Comparing lengths of these directed paths, we see that $m > b$, and then after cancellation we obtain $uhv = vq^{m-b}$. Hence uh is conjugate to some power of q, and since uh is a conjugate of p, we conclude that p is conjugate to some power of q.

Now suppose that s is a suffix of d and write $d = cs$. With t as before, we now obtain

$$\begin{aligned}(t,s)\,(d,pd)\,(s,t) &= (vq^b k, s)\,(cs, pcs)\,(s, vq^b k)\\ &= (cvq^b k, pcvq^b k) \in L_{q,k}.\end{aligned}$$

It follows that $pcvq^b k = cvq^m k$ for some $m \geqslant 0$. Again $m > b$ and after cancellation we obtain $pcv = cvq^{m-b}$. Here we see directly that p is conjugate to a power of q.

Now suppose that s is a suffix of d and write $d = cs$, and that t is a suffix of k and write $k = jt$. We now obtain

$$\begin{aligned}(t,s)\,(d,pd)\,(s,t) &= (t,s)\,(cs,pcs)\,(s,t)\\ &= (ct, pct) \in L_{q,k}.\end{aligned}$$

Since by assumption p is not the empty path, we have $ct = wq^a k$ and $pct = wq^b k$ for some suffix w of q and some $a, b \geqslant 0$. Again comparing lengths, we see that $b > a$. Writing $q = lw$ we obtain, after cancellation of ct, that $p = wq^b l$ and here $wq^b l$ is a conjugate of q^{b+1}.

Hence for each possibility of s, we deduce from (4.2.1) that p is conjugate to some power of q.

Using equation (4.2.2) we deduce similarly that q is conjugate to some power of p. Again comparing lengths, we conclude that p and q are conjugate. □

For the polycyclic monoids P_n $(n \geqslant 2)$, we obtain the classification of closed inverse submonoid up to conjugacy given in [16, Theorem 4.4] by applying Theorem 4.2.4 to the graph Γ with one vertex and n loops labelled $a_1, \ldots, a_n$.

For the case $n = 1$, we obtain the graph inverse semigroup $G(\Gamma) = B \cup \{0\}$, where B is the bicyclic monoid.

A proper closed inverse subsemigroup of $G(\Gamma)$ cannot contain $\{0\}$ and so is a proper closed inverse subsemigroup of B. By Lemma 3.3.5, such a subsemigroup is either $E(B)$ itself or is of finite chain type. Part (1) of Theorem 4.2.4, then shows that all closed inverse subsemigroup of finite chain type in B are then conjugate.

By [16, Theorem 4.3] (see Theorem 4.1.2), a closed inverse subsemigroup of B of cycle type consists of elements of the form

$$(qp^r, qp^s) \quad \text{for} \quad r, s \geqslant 0\,,$$

where $p = a^m$ and $q = a^d$ for some d with $0 \leqslant d \leqslant m-1$: that is elements of the form

$$(a^{rm+d}, a^{sm+d})\ .$$

It is therefore a closed inverse subsemigroup M_d as in Theorem 3.3.6, and since two paths a^m and a^n are conjugate circuits if and only if $m = n$, then M_d is conjugate to $M_{d'}$ if and only if $d = d'$.

Theorem 4.2.5. *Let L be a closed inverse subsemigroup of finite chain type in $G(\Gamma)$, with minimal element (w, w). Then L has infinite index in $G(\Gamma)$ if and only if there exists a non-empty directed circuit c in Γ and a (possibly empty) directed path g from some vertex v_0 of w to a vertex of c and with g having no edge in common with $c \cup w$.*

Proof. A coset representative of L has the form (q, u) where q is some suffix of w, and (q, u) have the same initial vertex. If L has infinite index, then there are infinitely many distinct choices for (q, u) and since Γ is finite, there must be a directed circuit in Γ as described, see Figure 29 below.

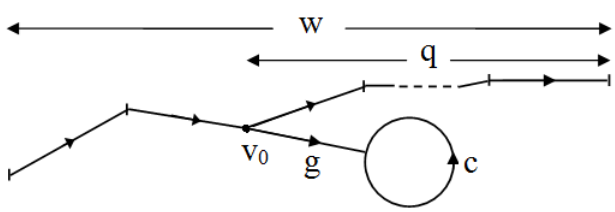

Figure 29: *Finite chain type.*

Conversely, suppose that g, c exist. Let q be the suffix of w that has initial vertex v_0. Then $q \in L$ and so for any $k \geqslant 0$ the coset $C_k = L(q, gc^k)^{\uparrow}$ exists.

Now for $k > l,$ we have if g is non-empty, that

$$(q, gc^k)(q, gc^l)^{-1} = (q, gc^k)(gc^l, q) = 0 \notin L$$

and so the cosets C_k and C_l are distinct. If g is empty then we have

$$(q, c^k)\,(q, c^l)^{-1} = (q, c^k)\,(c^l, q) = (q, c^{k-l}q) \notin L$$

and again the cosets C_k and C_l are distinct. □

Theorem 4.2.6. *Let L be a closed inverse subsemigroup of infinite chain type in $G(\Gamma)$. Then L has infinite index in $G(\Gamma)$.*

Proof. Here the elements of L consist of the idempotents determined by an infinite sequence of directed paths in $\Gamma,$ each of which is a suffix of the other. Eventually then, we find a path cq where c is a directed cycle, and $(cq, cq) \in L,$ as in Figure 30 below.

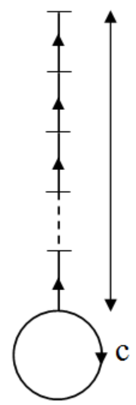

Figure 30: *Infinite chain type*

Then for each $k \geqslant 0,$ the element $(q, c^k q)$ represents a coset $C_k = L(q, c^k q)^{\uparrow}$.

Now for $k > l$,

$$(q, c^k q)\,(q, c^l q)^{-1} = (q, c^k q)\,(c^l q, q) = (q, c^{k-l} q) \notin L$$

and the cosets C_k and C_l are distinct. $\square$

Theorem 4.2.7. *If there is a closed inverse subsemigroup $L_{p,d}$ of cycle type in $G(\Gamma)$ such that p is a circuit with at least two distinct edges, then $L_{p,d}$ has infinite index in $G(\Gamma)$.*

Proof. Write $p = uv$ where each of u, v is non-empty and one contains an edge not in the other. Let c be the circuit vu (see Figure 31).

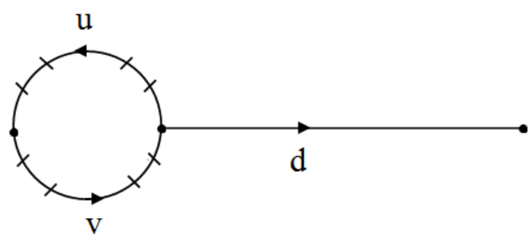

Figure 31: *Cycle type I.*

Then for $k \geqslant 1$, the element $(vd, c^k) \in G(\Gamma)$ and determines a coset $C_k = L_{p,d}(vd, c^k)^{\uparrow}$. Then for $k > l$,

$$(vd, c^k)\,(vd, c^l)^{-1} = (vd, c^k)\,(c^l, vd) = (vd, c^{k-l} vd) = (vd, up^{k-l} d) \notin L_{p,d}$$

and the cosets C_k and C_l are distinct. $\square$

We can now consider a graph Γ containing an edge a that is a directed circuit of length 1, and a closed inverse subsemigroup $L_{a^m, d}$ of cycle type.

Theorem 4.2.8. *A closed inverse subsemigroup $L = L_{a^m, d}$ of $G(\Gamma)$ is of infinite index if there exists a directed cycle c in Γ that contains an edge e with $a \neq e$, and a (possibly empty) directed path g from some vertex v_0 of d to a vertex of c and with g having no edge in common with $c \cup d$.*

Proof. Suppose that c, g exist and let q be the suffix of d with initial vertex v_0. Two cases with $c = ae$ are shown in Figures 32 and 33.

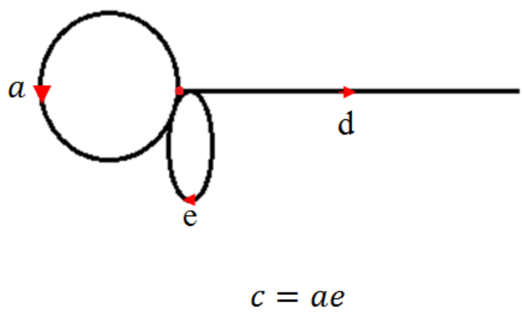

Figure 32: *Cycle type IIa.*

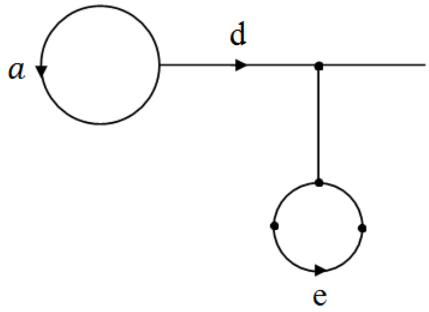

Figure 33: *Cycle type IIb.*

Let $C_k = L(q, gc^k)^{\uparrow}$. Then if g is non-empty, for $k > l$,

$$(q, gc^k)\,(q, gc^l)^{-1} = (q, gc^k)\,(gc^l, q) = 0 \notin L$$

and the cosets C_k and C_l are distinct. If g is empty, then

$$(q, c^k)\,(q, c^l)^{-1} = (q, c^k)\,(c^l, q) = (q, c^{k-l}q) \notin L$$

and the cosets C_k and C_l are again distinct. $\square$

We are now reduced to the case that the only directed circuits in Γ that can be attached to a vertex of d are powers of the loop a, as in Figure 34.

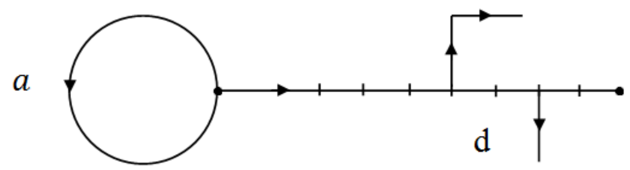

Figure 34: *Cycle type III.*

A coset representative of $L = L_{a^m,d}$ has the form $(a^r d, w)$ with $r \geqslant 0$, or (q, w) where q is a proper suffix of d. Hence w has the same initial vertex v_0 as d or of some proper suffix of d. We can only construct finitely many representatives of the form (q, w) and so we consider representatives of the form $(a^r d, w)$. Here w must have the form $w = a^s t$ for some $s \geqslant 0$ and some (possibly empty) directed path t with initial vertex v_0 and not containing the edge a. Fix t and consider the cosets $L(a^{r_1}d, a^{s_1}t)^{\uparrow}$ and $L(a^{r_2}d, a^{s_2}t)^{\uparrow}$ with $s_1 \geqslant s_2$. Now

$$(a^{r_1}d, a^{s_1}t)(a^{r_2}d, a^{s_2}t)^{-1} = (a^{r_1}d, a^{s_1}t)(a^{s_2}t, a^{r_2}d) = (a^{r_1}d, a^{s_1-s_2+r_2}d)$$

and $(a^{r_1}d, a^{s_1-s_2+r_2}d) \in L$ if and only if $r_2 - s_2 \equiv r_1 - s_1 \pmod{m}$. Hence for a fixed t we can produce exactly m distinct cosets of the form $L(a^r d, a^s t)^{\uparrow}$.

But for distinct paths t_1 and t_2, $a^{s_1}t_1$ cannot be suffix comparable with of $a^{s_2}t_2$, and so

$$(a^{r_1}d, a^{s_1}t_1)(a^{r_2}d, a^{s_2}t_2)^{-1} = 0 \neq L$$

and the cosets determined by distinct paths t_1 and t_2 are distinct.

Therefore we have:

Theorem 4.2.9. *Let $L = L_{a^m,d}$ where a is a directed circuit in Γ of length one, and there are no other directed circuits in Γ attached to a vertex of d. Then L has finite index in $G(\Gamma)$.*

Example 4.2.10. Suppose that Γ consists only of the loop a. Then the graph inverse semigroup $G(\Gamma)$ is the bicyclic monoid B (as in Chapter 3) with a zero adjoined:

$$G(\Gamma) = B \sqcup \{0\} .$$

A proper closed inverse submonoid of $G(\Gamma)$ is therefore a closed inverse submonoid of B, and the classification of Theorem 4.2.1 recovers that given in Lemma 3.3.5 and Theorem 3.3.6. From Theorems 4.2.5 and 4.2.6 we recover Theorem 3.3.8, the closed inverse submonoids contained in $E(B)$ have infinite index. Theorem 4.2.9 gives us the result in Theorem 3.3.9, that the closed inverse submonoid $M_d = L_{a^d}$ of B has index d.

Chapter 5

Finite Index and Finite Generation

The well-known Schreier Theorem asserts that a subgroup of finite index in a finitely generated group is finitely generated. Standard proofs using Cayley representation and actions of groups may be found in books such as [3, Corollary 3.1.1]. For a subgroup H of a free group $F,$ the Schreier Index Formula computes the rank of H from the rank of F and the index $[F : H]$. However, finitely generated subgroups of finitely generated groups need not be of finite index, even in the case of subgroups of free groups.

Margolis and Meakin [18, Theorem 3.7] proved that a closed inverse submonoid of a free inverse monoid is finitely generated if and only if it has finite index. They used methods based on relationships between immersions of graphs and closed inverse submonoids of free inverse monoids. Here, 'finitely generated' is interpreted as 'finitely generated *as a closed inverse submonoid*', so that the natural partial order is taken into account.

In this chapter, we shall look at the properties of closed inverse submonoids of inverse monoids considered in [18, Theorem 3.7], including being finitely generated (in the closed sense) and having finite index, and study their inter-relationships.

We will need some basic notions and results from the theory of groups and inverse monoids.

5.1 Preliminaries

Throughout this chapter, M will be an inverse monoid generated by a finite subset X. This means that the smallest inverse submonoid $\langle X \rangle$ of M that contains X is M itself: equivalently, each element of M can be written as a product of elements of X and inverses of elements of X. We write $A = X \cup X^{-1}$. Then each element of M can be written as a product of elements in A.

Definition An inverse monoid M is said to be *finitely generated in the closed sense* if there exists a finite subset $X \subseteq M$ such that, for each $m \in M$ there exists a product w of elements of X and their inverses such that $m \geqslant w$. Equivalently, the smallest closed inverse submonoid of M that contains X is M itself.

Clearly, any finitely generated inverse monoid is generated in the closed sense. The converse is not true: any inverse monoid with zero is generated in the closed sense by $\{0\}$ since, for all $m \in M$ we have $m \geqslant 0$. Another example is given by:

Example 5.1.1. Take any group G and any finitely generated subgroup K. Let

$$S = \bigcup_{i=1}^{\infty} G_i \quad (G_i = G)$$

$$H = \bigcup_{i=1}^{\infty} K_i \quad (K_i \leqslant G_i)$$

and make S an inverse semigroup as an infinite ascending chain of the $G_i's$ and H an inverse subsemigroup of S.

If $g \in G_i$ and $h \in G_j$, then in S,

$$g * h = g\, h \in G_{min\{i,j\}}$$

Semilattice of groups, with all homomorphisms $\phi_{ij} : G_i \longrightarrow G_j$ equal to id.

We define the natural partial order on S as follows: if $g \in G_k$ and $h \in G_l$, then

$$g \geqslant h \quad \textit{if and only if} \quad k \geqslant l \quad \textit{and} \quad g = h \;\; \textit{in } G \; .$$

Therefore, if $K_1 \leqslant G_1$ is the copy of K in G_1, then $H = (K_1)^{\uparrow}$, and hence H is finitely generated in the closed sense, but not finitely generated in the basic sense.

We remark that in [18] the notation $\langle X \rangle$ is used for the smallest *closed* inverse submonoid of M that contains X. We shall use $\langle X \rangle^{\uparrow}$ for this.

We will need to use some ideas from the theory of finite automata and for background information on this topic we refer to [15, 28].

Definition A *deterministic finite state automaton* $\mathcal{A}$ (or just an *automaton* in this chapter) consists of

- a finite set S of *states*,
- a finite *input alphabet* A,
- an *initial state* $s_0 \in S$,
- a partially defined *transition function* $\tau : S \times A \to S$,
- a subset $T \subseteq S$ of *final* states.

We shall write $s \lhd a$ for $\tau(s, a)$ if $\tau(s, a)$ is defined. Given a word $w = a_1 a_2 \cdots a_m \in A^*$ we write $s \lhd w$ for the state $(\ldots(s \lhd a_1) \lhd a_2) \lhd \cdots) \lhd a_m,$ that is, for the state obtained from s by computing the succesive outcomes of the transition function determined by the letters of w, with $s \lhd \varepsilon = s$ for all $s \in S$. We normally think of an automaton in terms of its *transition diagram*, in which the states are the vertices of a directed graph and the edge set is $S \times A$, with an edge (s, a) having $o(s, a) = s$ and $t(s, a) = s \lhd a$.

Each $a \in A$ determines a transformation $\tau_a : S \to S$ defined by $s\tau_a = s \lhd a$. Hence τ_a is a an element of the monoid $\mathcal{P}_S$ of all partially defined functions $S \to S$, and the functions

$a \mapsto \tau_a$ then induces a monoid homomorphism $\tau^* : A^* \to \mathcal{P}_S$, whose image is called the *transition monoid* of $\mathcal{A}$.

Let X be a finite set, X^{-1} a disjoint set of formal inverses of elements of X, and $A = X \sqcup X^{-1}$ An automaton $\mathcal{A}$ with input alphabet A is called a *dual automaton* if, whenever $s \lhd a = t$ then $t \lhd a^{-1} = s$. A dual automaton is called an *inverse automaton* if, for each $a \in A$ the partial function τ_a is injective. (See [14, Section 2.1].)

A word $w \in A^*$ is *accepted* or *recognized* by $\mathcal{A}$ if $s_0 \lhd w \in T$. The set of all words recognized by $\mathcal{A}$ is the *language* of $\mathcal{A}$:

$$L(\mathcal{A}) = \{w \in A^* : s_0 \lhd w \in T\}.$$

A language L is *recognizable* if it is the language recognized by some automaton. Let G be a group, with identity element 1_G, and suppose that G is finitely generated by a subset X. Define $A = X \sqcup X^{-1}$ and consider the induced monoid homomorphism $\theta : A^* \to G$. The inverse image of 1_G is the *word problem* of G. For a survey of some connections between group theory and formal language theory with emphasis on the word problem, see [30].

The connection between automata and closed inverse subsemigroups of finite index is made by the coset automaton.

Definition Let M be a finitely generated inverse monoid, generated by $X \subseteq M$, and let F be a closed inverse submonoid of M of finite index. Since M is generated by X, there is a natural monoid homomorphism $\theta : A^* \to M$. The *coset automaton* $\mathcal{C} = \mathcal{C}(M : F)$ is defined as follows:

- the set of states is the set of cosets of F in M,
- the input alphabet is $A = X \sqcup X^{-1}$,
- the initial state is the coset F,
- the transition function is defined by $\tau((Ft)^{\uparrow}, a) = (Ft(a\theta))^{\uparrow}$,
- the only final state is F.

By Lemma 2.4.2 and Proposition 2.4.3, $(Ft)^{\uparrow}\lhd a$ is defined if and only if $t(a\theta)(a\theta)^{-1}t^{-1} \in F$.

The following lemma is given in [18, Lemma 3.2] for the case that M is the free inverse monoid $\mathrm{FIM}(X)$.

Lemma 5.1.2. *The coset automaton of F in M is an inverse automaton. The language $L(\mathcal{C}(M : F))$ that it recognizes is*

$$F\theta^{-1} = \{w \in A^* : w\theta \in F\}$$

and $\mathcal{C}(M : F)$ is the minimal automaton recognizing $F\theta^{-1}$.

Proof. It follows from Proposition 2.4.3 that $\mathcal{C}(M : X)$ is inverse.

Suppose that w is recognized by $\mathcal{C}(M : F)$. Then $(w\theta)(w\theta)^{-1} = (ww^{-1})\theta \in F$ and $F(w\theta)^{\uparrow} = F$. From Proposition 2.2.5, we deduce that $w\theta \in F$.
Conversely, suppose that $w = a_{i_1} \dots a_{i_m} \in A^*$ and that $s = w\theta \in F$. For $1 \leqslant k \leqslant m$, write $p_k = a_{i_1} \dots a_{i_k}$, $q_k = a_{i_{k+1}} \dots a_{i_m}$, so that $w = p_k q_k$, and take $s_k = p_k\theta$, so that $s_1 = a_{i_1} \in A$. Then

$$s_1 s_1^{-1} s = s_1 s_1^{-1} s_1(q_2\theta) = s_1(q_2\theta) = w\theta = s$$

and so $s_1 s_1^{-1} \geqslant ss^{-1} \in F$. Therefore $s_1 s_1^{-1} \in F$ and $F \lhd a_{i_1} = (Fs_1)^{\uparrow}$ is defined. Now suppose that for some k we have that $F \lhd w_k$ is defined and is equal to $(Fs_k)^{\uparrow}$. Then

$$s_{k+1}\, s_{k+1}^{-1}\, s = s_{k+1}\, s_{k+1}^{-1}\, s_{k+1}(q_{k+1}\theta) = s_{k+1}\, (q_{k+1}\theta) = w\theta = s$$

and so $s_{k+1}\, s_{k+1}^{-1} \geqslant ss^{-1} \in F$ and therefore $s_{k+1}s_{k+1}^{-1} \in F$. But

$$s_{k+1}\, s_{k+1}^{-1} = s_k(a_{i_{k+1}}\theta)\,(a_{i_{k+1}}\theta)^{-1} s_k^{-1} \in F\,,$$

and so by Lemma 2.4.2, $(Fs_k)^{\uparrow}\lhd a_{i_{k+1}}$ is defined and is equal to $(Fs_k(a_{i_{k+1}}))^{\uparrow} = (Fs_{k+1})^{\uparrow}$. It follows by induction that $F \lhd w$ is defined in $\mathcal{C}(M : F)$ and is equal to $(Fs)^{\uparrow} = F$, and so $w \in L(\mathcal{C}(M : X))$. Now by a result of Reutenauer [23], a connected inverse automaton with one initial and one final state is minimal. $\square$

The set of *rational* subsets of a monoid M is the smallest collection that contains all the finite subsets of M and is closed under finite union, product, and generation of a submonoid. We recall the notion of *star-height* of a rational set from Chapter III of J. Berstel [5], *Transductions and Context-free Languages*. Version of December 14th 2009, online at www.infop7.org/file/659/LivreTransductions14dec2009.pdf.

Let M be a monoid. Define a sequence of subsets $\mathrm{Rat}_h(M)$, with *star-height* $h \geqslant 0$, recursively as follows:

$$\mathrm{Rat}_0(M) = \{X \subseteq M |\, X \text{ is finite}\},$$

and $\mathrm{Rat}_{h+1}(M)$ consists of the finite unions of sets of the form $B_1 B_2 \cdots B_m$ where each B_i is either a singleton or $B_i = C_i^*$, for some $C_i \in \mathrm{Rat}_h(M)$. It is well known that $\mathrm{Rat}(M) = \bigcup_{h \geqslant 0} \mathrm{Rat}_h(M)$.

A subset S of M is *recognizable* if there exists a finite monoid $N,$ a monoid homomorphism $\phi : M \to N$, and a subset $P \subseteq N$ such that $S = P\phi^{-1}$. We can describe this in another way using the syntactic congruence $\simeq_S$ on $M,$ defined as follows: for $a, b \in M$,

$$a \simeq_S b \iff \text{for all } m, n \in M \text{ we have } (man \in S \Longleftrightarrow mbn \in S)\,.$$

It is easy to see that $\simeq_S$ is a congruence on $M,$ and so the quotient set $M_S = M/\simeq_S$ is a monoid. Then S is recognizable if and only if M_S is finite: indeed, the quotient homomorphism $M \to M_S$ then recognizes S.

For free monoids $A^*,$ Kleene's Theorem (see for example [15, Theorem 5.2.1]) tells us that the rational and recognizable subsets coincide. For finitely generated monoids, we have the following theorem due to McKnight.

Theorem 5.1.3. *In a finitely generated monoid M, every recognizable subset is rational.*

If M is generated (as an inverse monoid) by X, then as above we have a monoid homomorphism $\theta : A^* \to M$. We say that a subset S of M is *recognized* by an automaton $\mathcal{A}$ if its full inverse image $S\theta^{-1}$ in A^* is recognized by $\mathcal{A}$.

Definition Given two words $u, v \in A^*$, an $L-$*extension* is a word z such that exactly one of the words uz, vz is in L. We say that $u \simeq_L v$ if no $L-$extension exists.

Theorem 5.1.4. *The language $L = \{w \in A^* : \ w\rho \in F\}$ is a regular language if and only if the equivalence relation $\simeq_L$ has finitely many classes.*

The motivation for the work in this chapter is the following theorem of Margolis and Meakin [18, Theorem 3.7].

Theorem 5.1.5. *Let $M = \mathrm{FIM}(X)$ be a free inverse monoid and N be a closed inverse submonoid of M. Then the following conditions are equivalent:*

(1) N is recognized by a finite inverse automaton,

(2) N has finite index in M,

(3) N is a recognizable subset of M,

(4) N is a rational subset of M,

(5) N is finitely generated in the closed sense.

The original theorem of [18, Theorem 3.7] includes an extra condition related to immersions of finite graphs, which we have omitted since it is only relevant to the case of $\mathrm{FIM}(X)$. We wish to understand the relationships between the conditions given in Theorem 5.1.5 for general inverse monoids M.

5.2 Finite Index implies Finite Generation

In this section, we consider a closed inverse submonoid F that has finite index in a finitely generated inverse monoid M. We shall show that F is finitely generated in the closed sense.

Our proof differs from that given in [18, Theorem 3.7] for the case $M = \mathrm{FIM}(X)$: instead we generalize the approach taken for groups in [3, Theorem 3.1.4]. Recall that a transversal to F in M is a choice of one representative element from each coset of F. We always choose the element 1_M from the coset F itself. For $s \in S$ we write $\overline{s}$ for the *representative* of the coset that contains s (if it exists).

Lemma 5.2.1. *Let R be a transversal to F in M and define, for $r \in R$ and $s \in M$, $\delta(r,s) = rs\,(\overline{rs})^{-1}$. Then*

(1) $(Fs)^{\uparrow} = (F\overline{s})^{\uparrow}$

(2) $\overline{\overline{s_1}\,s_2} = \overline{s_1\,s_2}$ *(where $s_1, s_2 \in M$)*

(3) $s = \delta(1_M, s)\,\overline{s}$.

Theorem 5.2.2. *A closed inverse submonoid of finite index in a finitely generated inverse monoid is finitely generated (in the closed sense).*

Proof. Let M be an inverse monoid generated by a set X. We may assume that if $x \in X$, then $x^{-1} \in X$. Therefore each $s \in M$ can be expressed as a product

$$s = x_1 x_2 \cdots x_n \,,$$

where $x_i \in X$. Suppose that F is a closed inverse subsemigroup of finite index in M. Let R be a transversal to F in M. Given $h \in F$, we write $h = x_1 x_2 \cdots x_n$ and consider the prefix $h_i = x_1 x_2 \cdots x_i$ for $1 \leqslant i \leqslant m$. Since

$$h_i h_i^{-1} h h^{-1} = h_i h_i^{-1} h_i x_{i+1} \cdots x_n h^{-1} = h_i x_{i+1} \cdots x_n h^{-1} = h h^{-1},$$

we have $h_i h_i^{-1} \geqslant h h^{-1}$. But $hh^{-1} \in F$ and F is closed, so that $h_i h_i^{-1} \in F$. Therefore the coset $(Fh_i)^{\uparrow}$ exists, and so does $\overline{h_i}$. Now consider the product

$$x_1 \cdot \overline{x_1}^{-1}\overline{x_1} \cdot x_2 \cdot \overline{x_1 x_2}^{-1}\overline{x_1 x_2} \cdot x_3 \cdot \overline{\overline{x_1 x_2} x_3}^{-1} \cdots \overline{x_1 x_2 \cdots x_{n-1}} \cdot x_n \leqslant x_1 x_2 \cdots x_n = h \,.$$

Note that $1_M = \overline{x_1 x_2 \cdots x_n}$.

Rewriting using part (2) of Lemma 5.2.1 we get

$$x_1 \cdot \overline{x_1}^{-1}\overline{x_1} \cdot x_2 \cdot \overline{\overline{x_1} x_2}^{-1}\overline{x_1 x_2} \cdot x_3 \cdot \overline{x_1 x_2 x_3}^{-1} \cdots \overline{x_1 x_2 \cdots x_{n-1}} \cdot x_n \leqslant x_1 x_2 \cdots x_n = h\,.$$

Now using the elements $\delta(r,s)$ from Lemma 5.2.1 we have

$$\delta(1_M, x_1)\delta(\overline{x_1}, x_2)\delta(\overline{x_1 x_2}, x_3) \cdots \delta(\overline{x_1 x_2 \cdots x_{n-1}}, x_n) \leqslant h\,.$$

To conclude the proof we show that $\delta(r,s) \in F$. But since $(Frs)^{\uparrow} = (F\overline{rs})^{\uparrow}$ then by Proposition 2.2.5(3) we deduce that $\delta(r,s) \in F$. $\square$

Corollary 5.2.3. *A closed inverse submonoid F of finite index r in a finitely generated inverse monoid M generated by n elements is finitely generated (in the closed sense) by at most nr elements.*

This is a fairly crude estimate, as seen in the next example.

Example 5.2.4. Let M be the Clifford inverse monoid $M = G_1 \cup G_0$ over the two-element semilattice $e_0 < e_1$ determined by a surjective homomorphism $\phi : G_1 \longrightarrow G_0$. The multiplication $\star$ in M is:

If	$a, b \in G_1$	then	$a \star b = ab \in G_1$,
If	$a, b \in G_0$	then	$a \star b = ab \in G_0$,
If	$a \in G_1, b \in G_0$	then	$a \star b = (a\phi)b \in G_0$,
If	$a \in G_0, b \in G_1$	then	$a \star b = a(b\phi) \in G_0$.

and the natural partial order on M is given by $g \geqslant g\phi$ for all g.

Suppose that G_1 can be generated by a subset X_1 of size k. Then clearly if e_0 is the identity element of G_0, M can be generated by the set $X = X_1 \cup \{e_0\}$ of size $k+1$. Now G_1 has index 1 in M but can be generated by fewer than $nr = k+1$ elements.

Example 5.2.5. In Example 5.2.4, we will take $H \leqslant G_0$ and $F = (H)^{\uparrow} = H \cup H\phi^{-1}$, where $H\phi^{-1} = \{g \in G_1 : g\phi \in H\}$, and F is a full closed inverse subsemigroup of G_0.

Choose $g \in G_1$ and then

$$\begin{aligned}(Fg)^{\uparrow} &= (H(g\phi) \cup (H\phi^{-1})g)^{\uparrow} \\ &= (H(g\phi))^{\uparrow} \cup (H\phi^{-1})g\ .\end{aligned}$$

Now

$$\begin{aligned}x \in H(g\phi)^{\uparrow} &\Leftrightarrow x \in H(g\phi) \text{ or } x\phi \in H(g\phi) \\ &\Leftrightarrow x \in (H(g\phi))\phi^{-1}.\end{aligned}$$

Therefore

$$(Fg)^{\uparrow} = H(g\phi) \cup (H(g\phi))\phi^{-1} \cup (H\phi^{-1})g.$$

But $(H\phi^{-1})g \subseteq (H(g\phi))\phi^{-1}$, and so

$$(Fg)^{\uparrow} = H(g\phi) \cup (H(g\phi))\phi^{-1}.$$

This implies that if H has finite index in G_0, then F has finite index in M. Moreover, if M is finitely generated then so is F.

On the other hand, consider $L \leqslant G_1$, then L is a closed inverse subsemigroup of M. A coset $(Ls)^{\uparrow}$ exists only if $ss^{-1} \in L$, only if $ss^{-1} = e_{G_1}$, only if $s \in G_1$. Write $s = g$, and so we have

$$(Lg)^{\uparrow} = Lg \quad \text{(standard coset in } G_1).$$

Hence L has finite index in $M \iff L$ has finite index in G_1 (in usual sense of groups). This implies that if G_1 is finitely generated (as a group), then so is L provided that it is of finite index.

The result below is a further consequence of Theorem 5.2.2.

Proposition 5.2.6. *Let M be a finitely generated inverse monoid. The intersection of a finitely generated closed inverse submonoid H of M and a closed inverse submonoid N of finite index in M is finitely generated (in the closed sense) in H.*

Proof. Let $x \in H \cap N$ and $m \geqslant x$: hence $m \in N$. Similarly $m \in H$ and so $H \cap N$ is closed in M.

Write $U = H \cap N$ and consider a coset $(Uh)^{\uparrow}$ with $h \in H$. Consider the mapping

$$(Uh)^{\uparrow} \mapsto (Nh)^{\uparrow}$$

from the set of cosets of U in M to the set of cosets of N in M. We have $hh^{-1} \in U = H \cap N$ and so $hh^{-1} \in N$. Hence the coset $(Nh)^{\uparrow}$ exists. Suppose that $(Uh)^{\uparrow} = (Uk)^{\uparrow}$. Then $hk^{-1} \in U$ and so $hk^{-1} \in N$. It follows that $(Nh)^{\uparrow} = (Nk)^{\uparrow}$ and hence the map is well-defined.

Now suppose that for $h, k \in H$ that $(Nh)^{\uparrow} = (Nk)^{\uparrow}$. Then $hk^{-1} \in N$. Clearly $hk^{-1} \in H$ so $hk^{-1} \in U$. Therefore $(Uh)^{\uparrow} = (Uk)^{\uparrow}$. So the set of cosets of U in H embeds into the set of cosets of N in M. So $H \cap N$ has finite index as in N and hence by Theorem 5.2.2, $H \cap N$ is finitely generated in the closed sense. $\square$

5.3 Recognizable Closed Inverse Submonoids

Theorem 5.3.1. *Let F be a closed inverse submonoid of a finitely generated inverse monoid M. Then the following are equivalent:*

(1) F is recognized by a finite inverse automaton,

(2) F has finite index in M,

(3) F is a recognizable subset of M.

Proof. If F has finite index in M, then by Lemma 5.1.2, its coset automaton $\mathcal{C}(M : F)$ is a finite inverse automaton that recognizes F. Conversely, suppose that $\mathcal{A}$ is a finite inverse automaton that recognizes F. Again by Lemma 5.1.2, the coset automaton $\mathcal{C}$ is minimal, and so must be finite. Hence (1) and (2) are equivalent.

If (2) holds, then we have a partial action of M on the set $\mathscr{C}$ of cosets of F giving, by Proposition 2.4.3 a homomorphism $M \to \mathscr{I}(\mathscr{C})$, in for which F is the inverse image of the stabilizer of the coset F. Therefore (2) implies (3).

To prove that (3) implies (2), suppose that F is recognizable and so the language

$$L = \{w \in A^* : \ w\rho \in F\}$$

is a regular language. By Nerode's Theorem 5.1.4, the equivalence relation $\simeq_L$ has finitely many classes. We claim that if $u \simeq_L v$ and if $(Fu)^\uparrow$ exists, then $(Fv)^\uparrow$ exists and $(Fu)^\uparrow = (Fv)^\uparrow$. Suppose that $u \simeq_L v$ and that $(F(u\rho))^\uparrow$ exists. So $(u\rho)\,(u\rho)^{-1} \in F$ and hence $(uu^{-1})\rho \in F$ and then $uu^{-1} \in L$. But by assumption $u \simeq_L v$, which means that no $L-$extension exists. It implies that $vu^{-1} \in L,$ and so by the definition of the language $L,$ we get $(v\rho)\,(u\rho)^{-1} \in F.$ Hence $(F(v\rho))^\uparrow = (F(u\rho))^\uparrow$. But by Nerode's Theorem 5.1.4, the relation $\simeq_L$ has finitely many classes and therefore there are finitely many cosets. It follows that F has finite index in $M.$ This proves the claim and hence (3) implies (2).

□

5.4 Rational Generation

In this section we give an automata-theory proof of a result of [18, Theorem 3.7]. Our proof is modeled on the approach in [8, Theorem II] to the proof of the following theorem due to Anisimov and Seifert [1, Theorem 3].

Theorem 5.4.1. *A subgroup of a finitely generated group G is a rational subset of G if and only if it is finitely generated.*

Theorem 5.4.2. *Let F be a closed inverse submonoid of a finitely generated inverse monoid M. Then N is generated (in the closed sense) by a rational subset if and only if F is generated (in the closed sense) by a finite subset.*

Proof. Since finite sets are rational sets, one half of the theorem is trivial.

So suppose that F is generated (in the closed sense) by some rational subset Y of F. If M is generated (as an inverse monoid) by X, we take $A = X \cup X^{-1}$, and as above let ρ be the obvious map $A^* \to M$. Then $Z = (Y \cup Y^{-1})^*$ is rational and then by the definition of rational sets, there exists a rational language R in A^* such that $R\rho = Z$.

The pumping lemma for R then tells us that there exists a constant C such that, if $w \in R$ with $|w| > C$, then $w = uvz$ with $|uv| \leqslant C$, $|v| \geqslant 1$, and $uv^i z \in R$ for all $i \geqslant 0$.

We set

$$U = \{uvu^{-1} : u, v \in A^*, |uv| \leqslant C, (uvu^{-1})\rho \in F\}$$

and $V = (U\rho)^\uparrow$ (the closed inverse submonoid of F generated by the image of U).

Clearly U is finite, and we claim that $F = V$.

First note that if $w \in R$ and $|w| \leqslant C$ then $w \in U$ (take $u = 1, v = w$) and so $w\rho \in V$. So 'short' words in R represent elements of V. But every element of F is above some element $w\rho$ in the natural partial order on F: what happens if w is a 'long' word?

Suppose that $|w| > C$ but that $w\rho$ has an element n of $F \setminus V$ above it. Choose $|w|$ minimal. The pumping lemma gives $w = uvz$ as above. Since $|uz| < |w|$ it follows that $(uz)\rho \in V$.

Moreover,

$$(uvu^{-1})\rho \geqslant (uvzz^{-1}u^{-1})\rho = (uvz)\rho((uz)\rho)^{-1} = (w\rho)\,((uz)\rho)^{-1}.$$

Now $w\rho \in F$ and $(uz)\rho \in V$: since F is closed, $(uvu^{-1})\rho \in F$ and therefore $uvu^{-1} \in U$. Now

$$n \geqslant w\rho = (uvz)\rho \geqslant (uvu^{-1}uz)\rho = (uvu^{-1})\rho\ (uz)\rho \in V.$$

Since V is closed, $n \in V$. But this is a contradiction. Hence $F = V$. □

Corollary 5.4.3. *If a closed inverse submonoid F of a finitely generated inverse monoid M is a rational subset of M then it is finitely generated (in the closed sense).*

Proof. If F is a rational subset of M then it is certainly generated by a rational set, namely F itself. □

However, the converse of Corollary 5.4.3 is not true. Before showing that, we prove the following lemma:

Lemma 5.4.4. *Let M be a semilattice of groups $G_1 \sqcup G_0$ over the semilattice $1 > 0$, and suppose that T is a rational subset of M of star-height h. Then $G_1 \cap T$ is a rational subset of G_1.*

Proof. We proceed by induction on h. If $h = 0$ then T is finite, and $G_1 \cap T$ is a finite subset of G_1 and so is a rational subset of G_1, also of star-height $h_1 = 0$.

If $h > 0$, then, as above, T is a finite union $T = S_1 \cup \cdots \cup S_k$ where each S_j is a product $S_j = R_1 R_2 \cdots R_{m_j}$ and where each R_i is either a singleton subset of M or $R_i = Q_i^*$ for some rational subset Q_i of M of star-height $h - 1$

Hence

$$G_1 \cap T = (G_1 \cap S_1) \cup \cdots \cup (G_1 \cap S_k)\,.$$

Consider the subset $G_1 \cap S_j = G_1 \cap R_1 R_2 \cdots R_{m_j}$. We claim that

$$G_1 \cap R_1 R_2 \cdots R_{m_j} = (G_1 \cap R_1)(G_1 \cap R_2) \cdots (G_1 \cap R_{m_j})\,. \tag{5.4.1}$$

The inclusion $\supseteq$ is clear, and so now we suppose that $g \in G_1$ is a product $g = r_1 r_2 \cdots r_{m_j}$ with $r_i \in R_i$. If any $r_i \in G_0$ then $g \in G_0$: hence each $r_i \in G_1$ and so $g \in (G_1 \cap R_1)(G_1 \cap R_2) \cdots (G_1 \cap R_{m_j})$, confirming (5.4.1).

The factors on the right of (5.4.1) are either singleton subsets of G_1, or are of the form $G_1 \cap Q_i^*$ where Q_i is a rational subset of M of star-height $h-1$. However, $G_1 \cap Q_i^* = (G_1 \cap Q_i)^*$: the inclusion $G_1 \cap Q_i^* \supseteq (G_1 \cap Q_i)^*$ is again obvious, and $G_1 \cap Q_i^* \subseteq (G_1 \cap Q_i)^*$ since if $w = x_1 \ldots x_m \in Q_i^*$ and some $x_j \in G_0$ then $w \in G_0$. It follows that if $w \in G_1 \cap Q_i^*$ then $x_j \in G_1$ for all j.

Hence $G_1 \cap T$ is a finite union of subsets, each of which is a finite product of singleton subsets of G_1 and subsets of the form $(G_1 \cap Q_i)^*$ where, by induction $G_1 \cap Q_i$ is a rational subset of G_1 of star-height $h_2 \leqslant h-1$. Therefore $G_1 \cap T$ is a rational subset of G_1 $\square$

Corollary 5.4.5. *Let $L = L_1 \sqcup L_0$ be an inverse submonoid of M that is also a rational subset of M. Then L_1 is a rational subset of M_1.*

Proof. Take $T = L$: then $G_1 \cap L = L_1$ is a rational subset of G_1. $\square$

Now, we show that the converse of Corollary 5.4.3 is not true.

Example 5.4.6. Let F_2 be a free group of rank 2 and consider the Clifford inverse monoid M (as in Example 5.2.4) determined by the abelianisation map $\alpha : F_2 \to F_2^{ab}$ where $F_2^{ab} \cong \mathbb{Z} \oplus \mathbb{Z}$. The kernel of α is the commutator subgroup F_2' of F_2 and we let K be the closed inverse submonoid $F_2' \sqcup \{\mathbf{0}\}$.

$$\begin{tikzcd} F_2' \arrow[r] \arrow[d] & F_2 \arrow[d, "\alpha"] \\ \{\mathbf{0}\} \arrow[r] & F_2^{ab} \end{tikzcd}$$

Now K is generated (in the closed sense) by $\{\mathbf{0}\}$ and so is finitely generated. But F_2' is not finitely generated (see, for example, [3, Example III.4(4)]) and so is not a rational subset of F_2 by the theorem of Anisimov and Seifert. Therefore using Corollary 5.4.5, K is not a rational subset of M.

This same example also gives us a counterexample to the converse of Theorem 5.2.2: K is finitely generated in the closed sense, but has infinite index in M.

5.5 Concordance

An inverse monoid M is *concordant* if every closed inverse submonoid F of M that is finitely generated (in the closed sense) and with $F \neq E(F)$ has finite index in M.

Example 5.5.1. Trivially, all finite inverse monoids are concordant, and the result of Margolis and Meakin [18, Theorem 3.7] shows that finitely generated free inverse monoids are concordant. The bicyclic monoid is concordant by Theorem 3.3.6 and Theorem 3.3.9.

Proposition 5.5.2. *A group G is concordant if and only if it is either finite or is infinite cyclic.*

Proof. Suppose that G is concordant. For each non-trivial element $g \in G$, the cyclic subgroup $\langle g \rangle$ has finite index in G. If one non-trivial element $g \in G$ has finite order, then G is finite. Otherwise, G is torsion-free, and has an infinite cyclic subgroup of finite index, and hence is itself infinite cyclic (see, for example, [29, Theorem 4.20]). □

The key idea in [18, Theorem 3.5] used to show that a finitely generated inverse submonoid of a free inverse monoid has finite index is the flower automaton. Suppose that M in an inverse monoid generated by the finite subset $X \subseteq M$. As before we write $A = X \sqcup X^{-1}$ and we have $\theta : A^* \to M$. Suppose that the closed inverse submonoid F is generated in the closed sense by $\{w_1\theta, \ldots, w_n\theta\}$ where each w_i is expressed as a word on the alphabet A. Let $Y = \{w_1, \ldots, w_n\} \subseteq A^*$. The *flower automaton* $\mathcal{F}(Y)$ is defined by its transition diagram, constructed as follows: We begin with a vertex v_0 and, for each $w_i \in Y$ we attach to v_0 a directed cycle – called a *petal* – labelled by w_i. The distinguished vertex v_0 is then both the initial state and the unique final state of $\mathcal{F}(Y)$.

Lemma 5.5.3. *There is a map of automata $\varphi : \mathcal{F}(Y) \to \mathcal{C}(M : F)$ defined as follows. We map $v_0 \mapsto F$. Every other state τ in $\mathcal{F}(Y)$ lies on a unique petal determined by $w \in Y$, and is the terminal vertex of some prefix p of w. Then $\tau\varphi = F \lhd p = F(p\theta)^{\uparrow}$. The transition described by the edge in $\mathcal{F}(Y)$, labelled by $a \in A$, from state τ to state $\tau \lhd a$ maps to the transition in $\mathcal{C}(M : F)$ from $F(p\theta)^{\uparrow}$ to $F((p\theta)a)^{\uparrow}$.*

Proof. The only thing that needs verifying is that the coset $F(p\theta)^{\uparrow}$ actually exists. But since p is a prefix of w, we have $w = pq$ for some q and so $(p\theta)(q\theta) = w\theta \in F$, and Lemma 1.3.7 applies. □

Lemma 5.5.4. *If $w \in Y$ determines a petal of the flower automaton and p is a prefix of w with $w = pq$ then $F(p\theta)^{\uparrow} = F(q^{-1}\theta)^{\uparrow}$. Hence to find $s\varphi$ for a state on a petal of $\mathcal{F}(Y)$, we can use a path either way around the petal, using the prefix p of w or the prefix q of w^{-1}.*

Proof. This is immediate from part (3) of Proposition 2.2.5, since $(p\theta)(q\theta) \in F$. □

The result of [18, Theorem 3.5] is then essentially the following:

Proposition 5.5.5. *If M is a free inverse monoid $M = \mathrm{FIM}(X)$, then the mapping*

$$\varphi : \mathcal{F}(Y) \to \mathcal{C}(\mathrm{FIM}(X) : F)$$

is surjective.

Proof. In this proof we shall not distinguish between a word $w \in A^*$ and its image in $\mathrm{FIM}(X)$. Consider a coset $(Fu)^{\uparrow}$ as a state of $\mathcal{C}(\mathrm{FIM}(X) : F)$ where $u \notin F$. The element $u \in \mathrm{FIM}(X)$ is represented by a Munn tree $(P, u\sigma)$ (see Section 1.11) , where σ is the minimum group congruence on $\mathrm{FIM}(X)$, and $\mathrm{FIM}(X)/\sigma = F(X)$. Then $(P, 1) = uu^{-1} \in F$, and so there exists some product $r = w_{i_1}^{\varepsilon_1} \cdots w_{i_k}^{\varepsilon_k}$ of elements of Y and their inverses with $r \leqslant uu^{-1}$. As a Munn tree, we have $r = (Q, 1)$ with $P \subseteq Q$.

We consider each w_j as a product of elements of $X \sqcup X^{-1}$. In constructing the connected subtree Q, each such element adds at most one edge to Q and so at some point we attach an edge that reaches the vertex $u\sigma \in P$ for the first time. At this stage we will have constructed a Munn tree $(R, u\sigma)$ with $R \subseteq Q$ and we will have used up some prefix $r'z$ of r, with

$$r' = w_{i_1}^{\varepsilon_1} \cdots w_{i_l}^{\varepsilon_l}$$

for some l with $0 \leqslant l \leqslant k$, and where z is either an empty string, a prefix of w_{l+1} or a prefix of w_{l+1}^{-1}, and $r'z\sigma = u\sigma$. Now

$$rr'z = (Q, 1)(R, u\sigma) = (Q \cup R, u\sigma) = (Q, u\sigma)$$

and

$$u(rr'z)^{-1} = (P, u\sigma)((u\sigma)^{-1}Q, u\sigma) = (P \cup Q, 1) = (Q, 1) = r \in F\,.$$

Therefore $(Fu)^{\uparrow} = (Frr'z)^{\uparrow}$. Since z is a prefix of an element of F, then by Lemma 1.3.7 we have $zz^{-1} \in F$.

Therefore $(Fz)^{\uparrow}$ exists, and since $rr'zz^{-1} \in F$, then $(Fz)^{\uparrow} = (Frr'z)^{\uparrow} = (Fu)^{\uparrow}$.

In the flower automaton r' determines a transition from v_0 completely around l petals and then either stops at state $\tau = v_0$ or follows a non-empty string partway around a leaf to stop at a state τ of $\mathcal{F}(Y)$. Then

$$\tau\varphi = (Fz)^{\uparrow} = (Fu)^{\uparrow}\,.$$

□

In general, the map $\varphi : \mathcal{F}(Y) \to \mathcal{C}(M : F)$ need not be surjective, as the following example shows.

Example 5.5.6. We return to Example 2.3.4 and take

$$Y = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & * & * \end{pmatrix} \right\} = \{y\}\,.$$

Then $F = \langle y \rangle^{\uparrow} = stab(1) = \{\sigma \in \mathscr{I}(\{1,2,3\}) : 1\sigma = 1\}$. The flower automaton $\mathcal{F}(Y)$ consists of the vertex v_0 and a single loop at v_0 labelled by y. However, as shown in Example 2.3.4, there are three cosets of F in $\mathscr{I}(\{1,2,3\})$ and so the coset automaton $\mathcal{C}(\mathscr{I}(\{1,2,3\}) : F)$ has three states.

We summarize our findings about the conditions considered by Margolis and Meakin in [18, Theorem 3.5].

Theorem 5.5.7. *Let F be a closed inverse submonoid of the finitely generated inverse monoid M and consider the following properties that F might have:*

(1) F is recognized by a finite inverse automaton,

(2) F has finite index in M,

(3) F is a recognizable subset of M,

(4) F is a rational subset of M,

(5) F is finitely generated in the closed sense.

Then properties (1), (2) and (3) are equivalent: each of them implies (4), and (4) implies (5). The latter two implications are not reversible.

Proof. The equivalence of (1), (2) and (3) was established in Theorem 5.3.1, and that (4) implies (5) in Corollary 5.4.3. The implication that (3) implies (4) is McKinight's Theorem 5.1.3.

Counterexamples for (5) implies (4) and (5) implies (2) are given in Example 5.4.6 $\square$

The equivalence of (2) and (5) for non-idempotent submonoids is the notion of concordance as defined in Section 5.5. It would be interesting to understand the class of concordant submonoids in other terms.

Bibliography

[1] **A.V. Anisimov and F.D. Seifert**, *Zur algebraischen Charakteristik der durch kontextfreie Sprachen definierten Gruppen.* Elektron. Informationsverarb. Kybernet 11 (1975) 695-702.

[2] **C.J. Ash and T.E. Hall**, *Inverse semigroups on graphs*. Semigroup Forum 11 (1975) 140-145.

[3] **G. Baumslag**, *Topics in Combinatorial Group Theory.* Lectures in Mathematics, ETH Zurich, Birkhauser Verlag (1993).

[4] **R. Baer**, *Zur Einführung des Scharbegriffs*. J. Reine Angew. Math. 160 (1929) 199-207.

[5] **J. Berstel**, *Transductions and context-free languages.* Teubner, Stuttgart (1979).

[6] **J. Certaine**, *The ternary operation* $(abc) = ab^{-1}c$ *of a group*. Bull. Amer. Math. Soc vol. 49, no. 12 (1943) 869-877.

[7] **L. Descalço and N. Ruškuc**, *Subsemigroups of the bicyclic monoid*. Internat. J. Algebra Comput. 15 (2005) 37-57.

[8] **C. Frougny, J. Sakarovitch and P. Schupp**, *Finiteness conditions on subgroups and formal language theory.* Proc. London Math. Soc. (3) 58 (1989) 74-88.

[9] **R. Gray, N. Ruškuc**, *Green index and finiteness conditions for semigroups*. J. Algebra 320 (2008) 3145-3164.

[10] **J.A. Green**, *On the structure of semigroups.* Annals of Math. 54 (1951) 163-172.

[11] **J.M. Howie**, *Fundamentals of Semigroup Theory.* London Math. Soc. Monographs, Oxford University Press (1995).

[12] **A. Jura**, *Coset enumeration in a finitely presented semigroup*. Canad. Math. Bull. 21 (1) (1978) 37-46.

[13] **M.V. Lawson**, *Coverings and embeddings of inverse semigroups*. Proc. Edinburgh Math. Soc. 36 (1993) 399-419.

[14] **M.V. Lawson**, *Inverse Semigroups.* World Scientific (1998).

[15] **M.V. Lawson**, *Finite Automata*. Chapman & Hall / CRC Press (2004).

[16] **M.V. Lawson**, *Primitive partial permutation representations of the polycyclic monoids and branching function systems*. Periodica Mathematica Hungarica 58(2), (2009) 189-207.

[17] **J.C. Meakin**, *An Invitation to inverse semigroups*. Proc. Conf. on (Ordered structures and algebra of computer languages), Hong Kong June, 1991, World Scientific (1993) 91-115.

[18] **S.W. Margolis, J.C. Meakin**, *Free inverse monoids and graph immersions*. International Journal of Algebra and Computations, World Scientific Publication Company 3(1993) 79-99.

[19] **W.D. Munn**, *Free inverse semigroups.* Proc. London Math. Soc. 30 (1974) 385-404.

[20] **M. Petrich**, *Inverse Semigroups*. John Wiley and Sons, New York (1984).

[21] **G.B. Preston**, *Inverse semigroups.* J. London Math. Soc. 29 (1954) 396-403.

[22] **N.R. Reilly**, *Embedding inverse semigroups in bisimple inverse semigroups*. Quart. J. Math. (Oxford) (2)16 (1965) 183-187.

[23] **C. Reutenauer**, *Une topologie du monoïde libre.* Semigroup Forum 18 (1979) 33-50.

[24] **J. Rose**, *A Course on Group Theory*. Cambridge University Press (1978).

[25] **N. Ruškuc, R.M. Thomas**, *Syntactic and Rees indices of subsemigroups.* J. Algebra 205 (2) (1998) 435-450.

[26] **B. Schein**, *Cosets in groups and semigroups. in Semigroups with applications, (eds J. M. Howie, W. D. Munn, and H. J. Weinert),* World Scientific, Singapore (1992) 205-221.

[27] **J.-P. Serre**, *Trees*. Springer-Verlag (1980).

[28] **M. Sipser**, *Introduction to the Theory of Computation*. PWS Publishing Company (1997).

[29] **G.C. Smith and O.M. Tabachnikova**, *Topics in Group Theory.* Springer Undergraduate Mathematics Series, Springer-Verlag (2000).

[30] **I.A. Stewart and R.M. Thomas**, *Formal languages and the word problem for groups.* In *Groups St. Andrews 1997 in Bath, II*, London Math. Soc. Lecture Note Ser. 261, Cambridge Univ. Press (1999) 689–700.

[31] **V.V. Wagner**, *Generalized groups.* Doklady Akademiĭ Nauk SSSR 84 (1952) 1119-1122 (Russian).